STARCH-BINDING DOMAIN-CONTAINING PROTEIN 1:

A NOVEL PARTICIPANT IN GLYCOGEN METABOLISM

Sixin Jiang

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the Department of Biochemistry & Molecular Biology
Indiana University
June 2011

Accepted by the Faculty of Indiana University, in partial
fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

Peter J. Roach, Ph.D., Chair

Doctoral Committee

_____

Anna A. DePaoli-Roach, Ph.D.

March 24, 2011

_____

Robert A. Harris, Ph.D.

_____

Nuria Morral, Ph.D.

DEDICATION

This work is dedicated to my parents, Zhanyan Jiang and Zongshu Liu, whose unconditional love and constant support have given me strength to face all the challenges in life.

This work is also dedicated to my grandparents, who always had faith in me.

.

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere gratitude toward my mentor, Dr. Peter Roach who has been supported me with knowledge, inspiration and valuable advice. He truly encouraged me and guided me with all his patience to be an independent scientist as well as a critical thinker. Throughout these years, he has taught me more than science and I appreciate it with all my heart.

I would like to thank my committee members, Dr. Anna DePaoli-Roach, Dr. Robert Harris and Dr. Nuria Morral, for all of their brilliant advice and insightful discussion about my projects as well as their warm encouragement. They helped me grow. I am greatly inspired by all their passion and scientific wisdom.

I would like to thank everyone who I worked with, current and past, in the Roach and DePaoli-Roach labs. Dr. Alexander Skurat, Lanmin Zhai, Dr. Wei Wang, Dr. Julia Degler, Dr. Vincent Tagliabracci, Dr. Jose Irimia, Cathy Meyer, Dyann Segvich, Chandra Karthik, Christopher Contreras, Punitee Garyali , Dr. Chiharu Nakai, Katrina Hughes, Jennifer Gleissner, Caron Peper, Dr. Gretchen Parker and Amanda McGuire. I am grateful to having them not only as my colleagues but also as my friends. It was a great pleasure working and spending time with all of them.

I would like to thank our neighbor and collaborator Dr. Clark Wells and everyone in his lab, Brigitte Heller, Bill Ranahan, Jacob Adler and Whitney Smith. They all

helped me through immunofluorescence microscopy and they made work fun. I would like to thank Dr. Keith Condon for helping me with histology. I would like to thank everyone in the Department office of Biochemistry and Molecular Biology.

And finally, I would like to thank my family and friends who have always been there for me.

ABSTRACT

Sixin Jiang

STARCH-BINDING DOMAIN-CONTAINING PROTEIN 1:

A NOVEL PARTICIPANT IN GLYCOGEN METABOLISM

Glycogen, a branched polymer of glucose, acts as an intracellular carbon and energy reserve in many tissues and cell types. The breakdown of glycogen by hormonally regulated degradation involving the coordinated action of glycogen phosphorylase and debranching enzyme has been well studied. However, the importance of lysosomal disposal of glycogen has been underscored by a glycogen storage disorder, Pompe disease. This disease destroys tissues by over-accumulating glycogen in lysosomes due to a genetic defect in the lysosomal acid α-glucosidase. Details of the intracellular trafficking of glycogen are not well understood. Starch-binding domain-containing protein 1 (Stbd1) is a protein of previously unknown function with predicted hydrophobic N-terminus and C-terminal CBM20 carbohydrate binding domain. The protein is highly expressed in the liver and muscle, the major repositories of glycogen. Stbd1 binds to glycogen *in vitro* and *in vivo* with a preference for less branched and more phosphorylated polysaccharides. In animal models, the protein level of Stbd1 correlates with the genetic depletion of glycogen. Endogenous Stbd1 is found in perinuclear compartments in cultured mouse and rat cells. When over-expressed in cells, Stbd1 accumulates and coincides with glycogen and GABARAPL1, the autophagy protein. They form enlarged perinuclear structures which are abolished by removing the hydrophobic N-terminus of Stbd1. Stbd1,

with point mutations in the CBM20 domain, retains the perinuclear localization but without concentration of glycogen in this compartment. In cells that are stably over-expressing glycogen synthase, glycogen exists as large perinuclear deposits, where Stbd1 can also be present. Removing glucose from the culture leads to a breakdown of the massive glycogen accumulation into numerous smaller and scattered deposits which are still positive for Stbd1. Furthermore, the autophagy protein GABARAPL1 co-immunoprecipates and co-localizes with Stbd1 when co-expressed in cells. Point mutation or deletion of the autophagy protein interacting region on Stbd1 eliminates the interaction and co-localization with GABARAPL1 but not the characteristic perinuclear distribution of Stbd1. We propose that Stbd1 is involved in glycogen metabolism. In particular, it participates in the vesicular transfer of glycogen to the lysosome with the recruitment of autophagy related proteins GABARAPL1 and/or GABARAP, as these vesicles mature prior to lysosomal fusion.

Peter J. Roach, Ph.D., Chair

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A/Ala	Alanine
ADP	Adenosine diphosphate
AGL	Amylo-1,6-glucosidase, 4-α-glucanotransferase
AIM	Atg8 family interacting motif
Ams	α-mannosidase
AMP	Adenosine monophosphate
AMPK	AMP activated protein kinase
Ape1	aminopeptidase 1
ATG	Autophagy-related genes
ATP	Adenosine triphosphate
BE	Branching enzyme
BNIP3	BCL2/adenovirus E1B 19 kDa protein-interacting protein 3
BNIP3L	BNIP3-like
BRCA1	Breast cancer 1
C/Cys	Cysteine
$CaCl_2$	Calcium chloride
CALCOCO2	Calcium binding and coiled-coil domain 2
cAMP	3'-5'-cyclic adenosine monophosphate
CAZy	Carbohydrate active enzymes
CBM	Carbohydrate-binding module
CGTase	Cyclodextrin glucanotransferase
$CHCl_3$	Chloroform
CHC	Clathrin heavy chain
CK-1	Casein kinase-1
CK-2	Casein kinase-2
CRT	Calreticulin
Cvt	cytoplasm-to-vacuole targeting pathway
D/Asp	Aspartic acid
Da	Dalton
DAB	3,3'-diaminobenzidine

DBE	Glycogen debranching enzyme
ddH_2O	Double-distilled water
DFCP1	double FYVE-domain containing protein 1
DNA	Deoxyribonucleic acid
DSP	Dual specificity phosphatase
DTT	Dithiothreitol
DYRK1A	Dual specificity tyrosine-phosphorylation-regulated kinase 1A
E/Glu	Glutamic acid
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethyleneglycol-O, O'-bis(2-aminoethyl)-N, N, N', N'-tetraacetic acid
EM	Electron microscopy
Epm2a	Epilepsy progressive myoclonus type 2a
Epm2b	Epilepsy progressive myoclonus type 2b
ER	Endoplasmic reticulum
F/Phe	Phenylalanine
F6P	Fructose-6-phosphate
F-1,6-BP	Fructose-1,6-bisphosphate
FIP200	Focal adhesion kinase family interacting protein of 200 KDa
G/Gly	Glycine
G1P	Glucose-1-phosphate
G6P	Glucose-6-phosphate
G6Pase	Glucose-6-phosphatase
G6PDH	Glucose-6-phosphate dehydrogenase
G6PT	Glucose-6-phosphate translocase
GA	Glucoamylase
GAA	Lysosomal acid-α-glucosidase
GABA	γ-Aminobutyric acid
GAGARAP	$(GABA)_A$ receptor-associated protein
GABARAPL1	GABARAP-like 1

GABARAPL2	GABARAP-like 2
GABRG2	$(GABA)_A$ receptor γ-2S subunit
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GEC1	Glandular epithelial cell protein 1
GH	Glycoside hydrolase
GLUT	Sodium independent glucose transporter
GN	Glycogenin
GPh	Glycogen phosphorylase
GS	Glycogen synthase
GSK3	Glycogen synthase kinase-3
GSD	Glycogen storage disease
GWD	Glucan, water dikinase
H_2O	Water
H_2O_2	Hydrogen peroxide
HCl	Hydrochloric acid
$HClO_4$	Perchloric acid
HCIP	High-confidence candidate interaction proteins
HK	Hexokinase
H_2SO_4	Sulfuric acid
HRP	Horseradish peroxidase
HSS	High speed supernatant
HSP	High speed pellet
IAA	Isoamyl alcohol
I/Ile	Isoleucine
IR	Insulin receptor
IRS	Insulin receptor substrate
kDa	Kilodalton
KH_2PO_4	Potassium dihydrogen phosphate
KOH	Potassium hydroxide
L/Leu	Leucine
LAMP	Lysosome-associated membrane protein

LB	Lysogeny broth
LC3/MAP1LC3	Microtubule-associated protein 1 light chain 3
LD	Lafora disease
LiAc	Lithium acetate
LiCl	Lithium chloride
LIR	LC3 interacting region
LSS	Low speed supernatant
LSP	Low speed pellet
MEF	Mouse embryonic fibroblast
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulfate
MGSKO	Muscle glycogen synthase knockout
NaCl	Sodium chloride
Na_2CO_3	Sodium carbonate
Na_2HPO_4	Sodium phosphate, dibasic
NaH_2PO_4	Sodium phosphate, monobasic
NADP	Nicotinamide adenine dinucleotide phosphate
NaOAc	Sodium acetate
NBR1	Next to BRCA1 gene 1 protein
NDP52	Nuclear dot protein of 52 kDa
NH_4HCO_3	Ammonium bicarbonate
NH_4OH	Ammonium hydroxide
$(NH_4)_2SO_4$	Ammonium Sulphate
NHLRC1	NHL repeat-containing protein 1
NIP3	BNIP3
Nix	NIP3-like protein X
NMR	Nuclear magnetic resonance
NSF	N-ethylmaleimide sensitive fusion protein
OH	Hydroxyl
PAS	Periodic acid-Shiff; Phagophore-assembly site or pre-autophagosome structure

PCR	Polymerase chain reaction
PDK-1	Phosphoinositide-dependent kinase-1
PE	phosphatidylethanolamine
PEG	Polyethylene glycol
PFK	Phosphofructokinase
PGI	Phosphoglucose isomerase
PGM	Phosphoglucomutase
Ph	Phosphorylase
PhK	Phosphorylase kinase
PI(3)P	Phosphatidylinositol-3-phosphate
PI(4,5)P2	Phosphatidylinositol-(4,5)-bisphosphate
PI(3,4,5)P3	Phosphatidylinositol-(3,4,5)-triphosphate
PKA	cAMP dependent protein kinase / protein kinase A
PI3K	Phosphatidylinositol-3-kinase
PKB/Akt	Protein kinase B
PM	Plasma membrane
PMSF	Phenylmethylsulfonylfluoride
PP1c	Catalytic subunit of protein phosphatase 1
PP1G	Glycogen associated protein phosphatase 1
PTG	Protein targeting to glycogen
R_{GL}/G_M	Regulatory subunit of protein phosphotase 1, muscle isoform
RNA	Ribonucleic acid
RT	Room temperature
RT-PCR	Reverse transcription polymerase chain reaction
S/Ser	Serine
SBD	Starch-binding domain
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophresis
SEX4	Starch excess 4 protein
SM	Skeletal muscle
SQSTM1	Sequestosome protein-1, p62

SR	Sarcoplasmic reticulum
Stbd1	Starch-binding domain-containing protein 1, genethonin 1
T/Thr	Threonine
TBS	Tris-buffered saline
TBST	Tris-buffered salin containing tween-20
TCA	Trichloroacetic acid
TLCK	*N*-*p*-tosyl-l-lysine chloromethyl ketone
Tris	Tris(hydroxymethyl)aminomethane
UBL	Ubiquitin like
UDP	Uridine diphosphate
UDP-glucose	Uridine diphosphate glucose
ULK1	Unc-51 like kinase 1 (Atg1 ortholog)
UMP	Uridine monophosphate
UTP	Uridine triphosphate
W/Trp	Tryptophan
WT	Wild type
Y/Tyr	Tyrosine

INTRODUCTION

1. Glycogen and its metabolism

1.1 Glycogen structure and function

Glucose polymers are universally present in organisms from single cell microorganisms to humans, in the form of starch or glycogen. The widely accepted role of glycogen is as an intracellular carbon and energy reserve. As branched polymeric glucose, glycogen acts as a repository of the monosaccharide in many tissues and cell types [1, 2]. For whole body glucose homeostasis, the major storage of glycogen is in the liver and skeletal muscle but many other tissues are capable of synthesizing glycogen, such as heart, brain, kidney and adipose tissue. In all tissues, glycogen serves as an energy reserve, providing fuel for muscular activity and in liver supplying glucose for export to the bloodstream to prevent low blood glucose. In addition to its general metabolic function, glycogen has also been proposed to have other, more controversial roles. For instance, glycogen was reported to be relevant in chromatin condensation and nuclus formation in *Xenopus* eggs [3]. However, in *Saccharomyces cerevisiae*, cell division was not impaired by lacking glycogen [4, 5].

There are two types of glycosidic linkages between glucose units in glycogen. The formation of α-1,4-glycosidic linkages is required for polymer elongation, while the branch points are introduced by α-1,6-glycosidic linkages, at the C6-OH of one of every 8-12 glucose residues on average. The frequency and distribution of branches determines glycogen topology, structure and solubility, and differentiates glycogen from amylopectin, the carbohydrate moiety of starch [6]. The Meyer-Bernfeld model of the branching structure of glycogen has been generally accepted and refined [7] (Figure 1). In this model, a full size glycogen molecule is predicted to contain 55,000 glucose residues forming 12 tiers with a

diameter of ~40 nm and Mr ~10^7 [8]. These glucose chains can be classed as the outer A chains without branches and the inner B chains with two branch points. The number and length of A chains and B chains are roughly equal, normally 13 glucose residues in each chain.

Although there has been no experimentally determined three-dimensional structure, glycogen particles isolated from liver are classified in three different structure groups by electron microscopy (EM) [9]. The α-particles, with the typical rosette shape and relative molecular mass (Mr) ~10^8, are commonly present in liver. The smaller globular β-particles (Mr ~10^7) found in muscle and are ~20-30 nm in diameter and considered to be subunits of α-particles. The γ-particles are even smaller, described as 3 nm-subunits of α and β-particles.

Besides glucose, glycogen also contains minor components like glucosamine [10] and phosphate [11, 12]. Small amounts of glucosamine were found in rabbit and pig liver glycogen, but not in skeletal muscle and heart glycogen [10]. UDP-glucosamine is considered to be the glycogen synthase substrate that incorporates glucosamine into glycogen [13]. Liver glycogen purified from rats injected with galactosamine showed as much as 10% glucose residues replaced by glucosamine in α-1,4-linkages [10]. Covalently bound phosphate has been detected in rabbit skeletal muscle glycogen [11, 12]. These phosphates were suggested to be present as a phosphomonoester at the C6 and a phosphodiester, which introduces an alternative branch point between C1-C6 linkages [14]. A specific enzyme, UDP-glucose: glycogen glucose-1-phosphotransferase was proposed to form the diester bond using phosphate from UDP-glucose. However, the enzyme was not completely defined at the molecular level.

So far, the physiological significance of these atypical, covalent structural components of glycogen has not been fully understood. Recent studies in our laboratory proposed that hyperphosphorylation leads to abnormal glycogen structure [15], pointing to a critical role of covalent phosphate in maintaining

glycogen structure and solubility. Furthermore, glycogen synthase has been shown to incorporate the β-phosphate of UDP-glucose into glycogen during synthesis forming covalent phosphate linkage at the C2 and C3 of glucose in mammalian glycogen [16]. Therefore, a glycogen damage/repair process has been proposed in which phosphates are introduced by error and removed by subsequent processing [16].

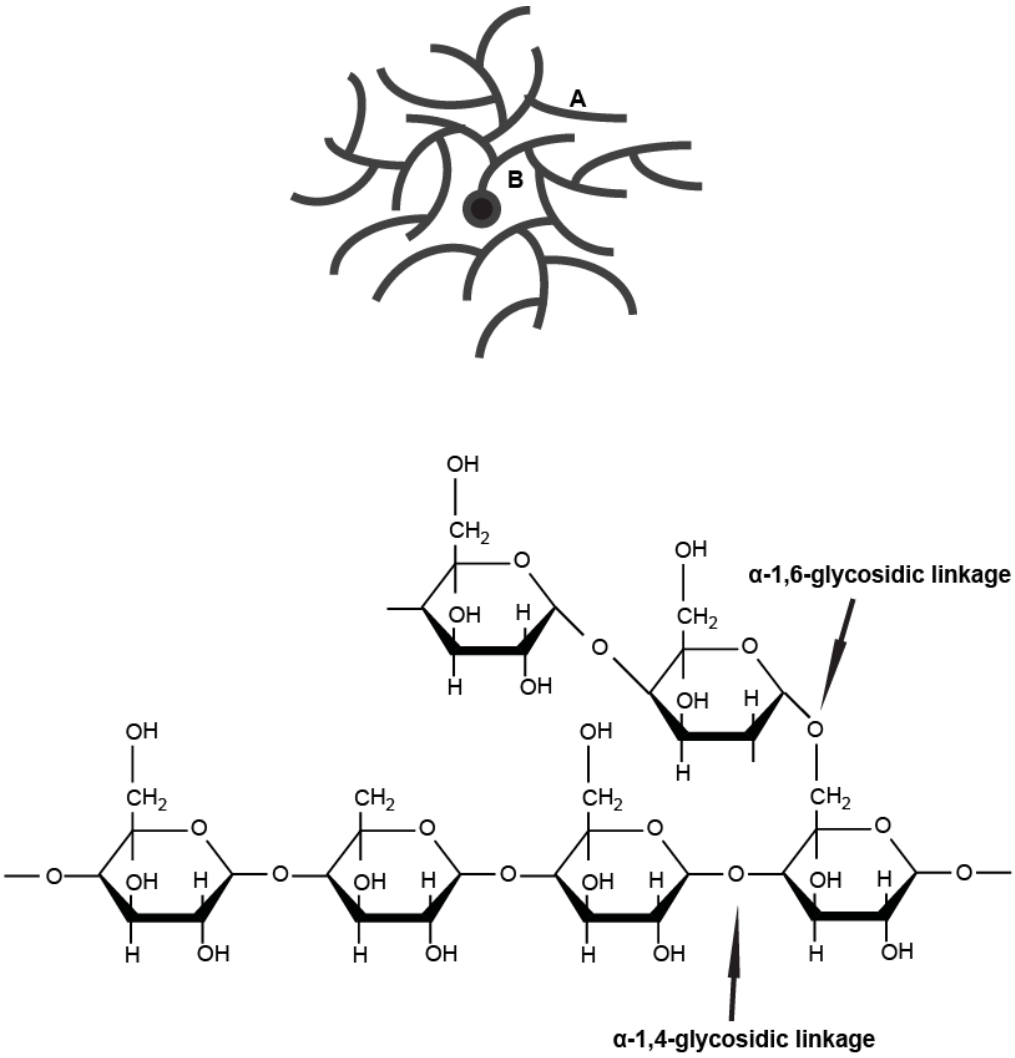


Figure 1. Glycogen structure and glycosidic linkages.

Glycogen is shown in the upper panel in branched pattern with outer unbranched A chains and inner B chains containing two branch points. The two types of glycosidic linkages are shown in the lower panel.

1.2 Subcellular distribution of glycogen

Glycogen is mostly described as cytosolic but glycogen particles have been found by electronic microscopy close to intracellular membranes such as endoplasmic reticulum (ER) in liver [17] and sarcoplasmic reticulum (SR) in muscle [17, 18]. Glycogen is also present in the yeast vacuole [19] and the lysosomes of mammalian cells. For example, about 10% glycogen was lysosomal in normal mammalian hepatocytes [20, 21]. This subcellular localization of glycogen would allow for a more complex metabolism and would suggest the presence of multiple glycogen pools. The α-glucosidase activity in the lysozome is critical for glycogen degradation as well as glycogen phosphorylase action in cytosol [1].

1.3 Glycogen metabolism

Generally, glycogen metabolism is mainly regulated by the coordination of the rate limiting enzymes of glycogenesis and glycogenolysis, i.e. glycogen synthase and glycogen phosphorylase, respectively. An overview of glycogen metabolism is shown in Figure 2. Glycogen synthesis is mediated by the combined actions of three enzymes, glycogenin (GN), glycogen synthase (GS) and the branching enzyme (BE). The cytosolic degradation of glycogen which can be hormonal regulated and exercise mediated, requires the coordinated action of glycogen phosphorylase (GPh) and the debranching enzyme (DBE/AGL), while the acid-α-glucosidase (GAA) is essential for glycogen hydrolysis in lysosomes.

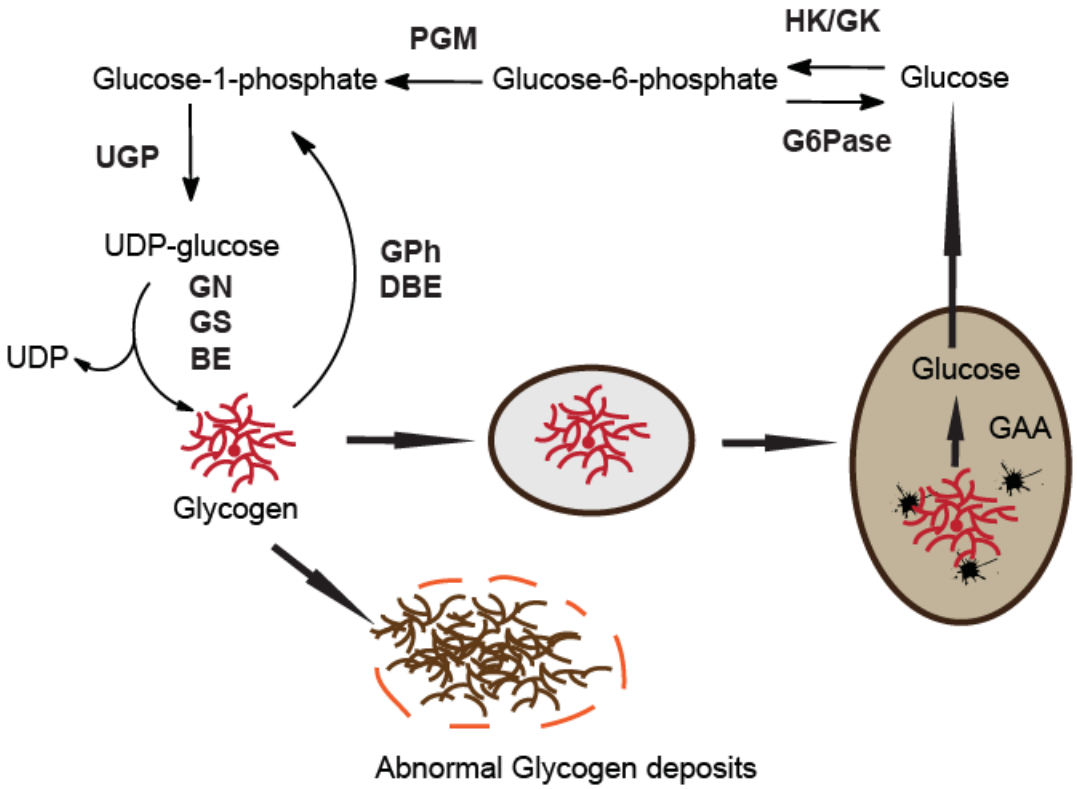


Figure 2. Schematic model of Glycogen metabolism.

HK/GK, hexokinase, glucokinase; PGM, phosphoglucomutase; G6Pase, glucose-6-phosphatase; UGP, UDP-glucose pyrophosphorylase; GN, glycogenin; GS, glycogen synthase; BE, branching enzyme; GPh, glycogen phosphorylase; DBE, debranching enzyme; GAA, lysosomal acid α-glucosidase.

1.3.1 Glycogen synthesis

After a meal, glucose is taken up by tissues and primarily converted to glycogen in skeletal muscle and liver. This helps keep blood glucose levels from being excessively elevated.

There are multiple glucose transporters in mammals that facilitate glucose entry into cells [22, 23]. They are different in regulation mechanisms and tissue distribution. Many tissues constitutively express GLUT1 at a basal level. In gluconeogenic tissues like liver and kidney, and in β-cells in the pancreas, glucose is taken into cells by the constitutively active GLUT2, which is independent of insulin regulation. In insulin-sensitive peripheral tissues, such as muscle and adipose tissue, GLUT4 is the dominantly expressed isoform, which translocates from an intracellular localization to the plasma membrane upon insulin stimulation [24].

Immediate conversion of intracellular glucose to glucose-6-phosphate (G6P) is mediated by hexokinase in muscle and glucokinase in the liver. When abundant nutrients are present, the glucose-6-P is converted to glucose-1-phosphate (G1P) by phosphoglucomutase (PGM). Uridine diphosphate glucose (UDP-glucose), which serves as the substrate for glycogen synthesis, is then formed from G1P and uridine 5'-triphosphate (UTP) by the action of UDP-glucose pyrophosphorylase (UGP).

In the 1930's, glycogen synthesis was thought to be the reversal of glycogen degradation by phosphorylase using G1P as substrate [25]. However, a separate synthetic enzyme, glycogen synthase, was discovered in the 1950's [26]. Studies of glycogen synthase suggested that a separate protein primer was required for glycogen synthesis, which led to the discovery of glycogenin in the 1970's.

1.3.1.1 Glycogenin (GN)

Glycogenin (EC2.4.1.186), the unique self-glycosylating protein, belongs to the family 8 retaining glycosyltransferases (GT8). It was found as the initiator of glycogen biosynthesis in the 1970's [27, 28] and acts as catalyst, substrate and product at the same time. The crystal structure of rabbit muscle glycogenin showed it to be a functional dimer of a 37 KDa subunit [29]. In humans, there are two isoforms of glycogenin. Glycogenin-1, encoded by the *GYG1* gene, is widely expressed especially in muscle. A mutation in *GYG1* gene that leads to the inactivation of self-glycosylation has been reported to cause glycogen depletion in skeletal muscle and accumulation of Periodic acid-Shiff (PAS)-positive material in cardiomyocytes leading to cardiomyopathy and cardiac arrhythmia [30]. Glycogenin-2, encoded by *GYG2* is limited to express in liver, cardiac muscle and pancreas [31].

GN initiates the synthesis of glycogen through a series of self-glucosylations using UDP-glucose as substrate to form a short, ~10-residue glucose chain. A C1-O-tyrosyl linkage is first formed at the tyrosine Y194 of GN and followed by the repeated formation of α-1,4-glycosidic linkages [32, 33]. The glucose oligomer then serves as a primer for the subsequent bulk glycogen synthesis [34, 35].

1.3.1.2 Glycogen synthase (GS)

As a member of the family 3 retaining glycosyltransferases (GT3), glycogen synthase (EC2.4.1.11) is the rate limiting intracellular enzyme in glycogen biosynthesis. It transfers the glucosyl group from UDP-glucose to the non-reducing end of a glycogen molecule and releases UDP, thereby forming the α-1,4-glycosidic linkages in glycogen (Figure 1 and 2). In mammals, two isoforms of GS, encoded by the *GYS1* and *GYS2* genes, have been identified with 70% sequence identity [1]. One isoform is expressed in muscle and other tissues

while the other is liver specific. Yeast GS also exists as two isoforms, 50% identical to mammalian sequences. Plant starch synthetases and bacterial starch/glycogen synthases are distinguished from mammalian glycogen synthases in two aspects. They use ADP-glucose instead of UDP-glucose as the glucosyl donor and they are not sensitive to G6P activation. Although there is not much overall sequence similarity, two sequence motifs were found in common [36].

GS activity is regulated by both reversible phosphorylation [37] and allosteric ligand binding, most importantly G6P. In general, GS is inactivated by phosphorylation via protein kinases and the potent allosteric activator G6P overcomes this inactivation. The ratio of the GS activities in the absence and presence of G6P can be used as a kinetic index of activation state [38]. The mammalian GS has two phosphorylation sites at the N-terminus and seven or five at the C-terminus, respectively, in the muscle and liver isoforms [1]. Various protein kinases has been reported to phosphorylate GS *in vitro*, such as glycogen synthase kinase-3 (GSK-3) [39, 40], phosphorylase kinase (PhK) [41], cAMP dependent protein kinase (PKA) [42], casein kinase 1 (CK1) [43, 44], casein kinase 2 (CK2) [45], AMP activated protein kinase (AMPK) [46], PAS domain-containing serine/threonine-protein kinase (PAS kinase) [47], dual specificity tyrosine-phosphorylation-regulated kinase 1A (DYRK1A) [48] and p38 MAPK [49].

When the inhibitory phosphates are removed from GS by glycogen-associated protein phosphatases1 (PP1Gs), GS is in an active state. PP1Gs are composed of the catalytic subunit of protein phosphatase 1 (PP1c) and a regulatory or targeting subunit that binds glycogen by carbohydrate binding domains, which lock the phosphatase to glycogen to dephosphorylate enzymes participating in glycogen metabolism [50]. To date, four of the regulatory subunits, R_{GL}, G_L, PTG and PPP1R6 have been well studied, and three new putative sequences (*PPP1R3E*, *PPP1R3F* and *PPP1R3G*) have been identified by bioinformatic

approaches [51]. Besides a carbohydrate binding domain, they all have a PP1 binding motif (RVXF) [52]. R_{GL} (also called G_M, encoded by *PPP1R3A*) was the first characterized glycogen-binding subunit [53]. The R_{GL} subunit is restricted to striated muscle and is critical for GS activation in response to exercise [54]. A prevalent mutation in *PPP1R3A* impairs glycogen synthesis, thus decreasing glycogen levels in skeletal muscle of human and mouse [55]. G_L (encoded by *PPP1R3B*), smaller than R_{GL}, is primarily expressed in liver and skeletal muscle in humans but only in liver of rodents [56]. G_L expression depends on insulin regulation and PP1-G_L activity is allosterically controlled by glucose and phosphorylase [57, 58]. PTG (also called PPP1R5 or R5, encoded by *PPP1R3C*) has been reported to be expressed in various tissues and interacts with glycogen metabolizing enzymes like GS, GPh and PhK [59-61] . PPP1R6 (R6, encoded by *PPP1R3D*), is widely expressed in tissues, with the highest levels in heart and skeletal muscle and much lower levels in the liver [62]. PPP1R3E (R3E, encoded by *PPP1R3E*) mRNA distribution is different in rat and human tissues. It is most abundant in rat liver and cardiac muscle, but barely detectable in human liver. This species-specific difference is similar to G_L mRNA, which is highly expressed in human but not rodent skeletal muscle. The activity of PP1-R3E was proposed to be under long-term control by insulin in rat liver [63].

The activated GS elongates the glucose chain which would eventually become insoluble and perhaps toxic to the cell. The formation of branch points is necessary to keep the polymer soluble, thus requiring another important glycogen metabolic enzyme, the branching enzyme.

1.3.1.3 The branching enzyme (BE)

The branching enzyme (EC2.4.2.18), with amylo-1, 4 to 1, 6-transglucosidase activity, maintains the structure and solubility of growing glycogen particles by introducing polysaccharide α-1,6-glycosidic linkages as branch points. Specifically, the branching enzyme cleaves an α-1,4-glycosidic linkage, cuts off a

section of glucosyl units and reforms an α-1,6-bond [64]. The mammalian BE works as a monomer, ~77 KDa [65] and shows sequence similarity with other polysaccharide enzymes that modify branched structures, e.g. α-amylase. The solubility of polysaccharides is determined by the branching frequency. Glycogen, branching is more frequent than amylopectin, making glycogen more soluble in aqueous solution. Therefore, one important function of glycogen branching is to keep the glucose polymer functional and non-toxic to the cells. No regulation of the BE has yet been established.

1.3.2 Glycogen degradation

When energy is necessary, in times of fasting or exercise, glycogen is broken down to release glucose for metabolism. Glycogen degradation happens in two ways. One is the hormonal or exercise mediated pathway which requires glycogen phosphorylase and debranching enzyme [1]. The other process involves transfer of glycogen to the lysosome and the action of lysosomal acid α-glucosidase (also named acid maltase). The significance of the lysosomal hydrolysis pathway is underlined by the fact that defects in this glucosidase cause a severe glycogen storage disease, Pompe disease, which eventually destroys tissues by over-accumulating glycogen in lysosomes [2, 66]. The mechanism of glycogen delivery to the lysosome is not well understood, and may involve some vesicular trafficking mechanism such as occurs in autophagy [67, 68].

1.3.2.1 Glycogen phosphorylase (GPh)

Glycogen phosphorylase (EC 2.4.1.1) (GPh) is affiliated with the family 35 glycosyltransferases. As one of the key enzymes in glycogenolysis, GPh drives the generation of glucose-1-phosphate and glycogen$_{(n-1)}$ by phosphorolytion and subsequent cleavage of the α-1,4-glycosidic linkages of glycogen (Figure 2). In mammals, three isoforms of GPh have been defined, in skeletal muscle, liver and

brain [69]. Since the three dimensional structure of the mammalian muscle isoform has been solved, studies are able to correlate catalysis and regulatory mechanisms with protein structures [70]. The most active form of the enzyme is considered to be a homodimer of the 97 KDa subunit [71] which is able to transition between an active R state and a less active T state by phosphorylation and allosteric ligand interaction. The phosphorylation triggers a conformational change so that the enzyme switches from the less active T state to the more active R state. The more active form of the enzyme, phosphorylase a, is phosphorylated by PhK at a single serine residue, S14. The protein phosphatase that is responsible for GPh dephosphorylation, much as described for GS dephosphorylation, is the PP1G, containing the R_{GL}, PTG-R5 or R6 regulatory subunits in skeletal muscle and PTG-R5, R6 or G_L in liver [72]. The allosteric activator AMP can restore activity to the dephosphorylated form of the enzyme, phosphorylase b, shifting back to the R state. Negative allosteric effectors like G6P and ATP compete with binding sites for AMP in the muscle enzyme. Phosphorylase activity is measured in the absence and presence of its allosteric regulator AMP similar to that for glycogen synthase. The resulting activity ratio is used as a kinetic index of enzyme activation on account of its covalent phosphorylation status [73].

1.3.2.2 The debranching enzyme (DBE)

The debranching enzyme (DBE, also called amylo-1,6-glucosidase, 4-α-glucanotransferase, AGL) (EC 3.2.1.68) is a monomer, ~165 KDa [74]. When phosphorylase stalls four glucose residues from an α-1,6-brachpoint, the DBE is required for the continuation of glycogen degradation. DBE contains two different active sites responsible for two individual catalytic activities [75]. The N-terminal domain is an oligo-1,4-1,4-glucantransferase, in the family 13 of glycoside hydrolases (GHs) [76] with sequence similarity to α-amylase. The N-terminal domain hydrolyzes an α-1,4-glycosidic bound and links the detached glucose residues to the end of the main chain by a reformed α-1,4-linkage,

leaving a single branched residue at the α-1,6-branchpoint. The C-terminal domain has amylo-1,6-glucosidase activity [75] that removes the remaining branched glucose residue.

1.3.2.3 Acid-α-glucosidase (GAA)

The acid-α-glucosidase (GAA) (EC 3.2.1.20), belongs to the family 31 glycosylhydrolases and breaks down glycogen into glucose by hydrolyzing both α-1,4- and α-1,6-linkages. The 76 kDa mature and active enzyme is generated from a 110 kDa precursor via proteolysis at both the N- and the C-termini [77]. As a glycoprotein with seven sites of N-glycosylation [78], GAA shares significant sequence similarity to sucrase-isomaltase, a bifunctional enzyme and an integral membrane protein in the small intestine brush border membrane [79]. Some catalytically critical residues have been identified in a region of conserved sequence among α-glucosidase, sucrase and isomaltase catalytic domains [80]. Transcriptional regulation of the human GAA gene has been reported [81], but the enzyme activity seems mostly to depend on pH.

1.4 Regulation of glycogen metabolism

Skeletal muscle and liver have the majority of glycogen deposits that are pivotal in whole body glucose homeostasis.

1.4.1 Regulation in skeletal muscle

As an insulin sensitive tissue, the conversion of glucose in skeletal muscle is largely regulated by insulin. Insulin promotes glucose uptake by regulating GLUT4 translocation and stimulates phosphatidylinositol-3-kinase (PI3K) mediated signaling. Increased glucose uptake results in an elevation of G6P levels, which activates glycogen synthase allosterically to stimulate glycogen synthesis. In the signaling cascades, insulin first binds to the extracellular α-

subunit of the insulin receptor (IR) to increase its tyrosine kinase activity. Then a series of tyrosine residues in the intracellular β-subunit of the IR undergoes autophosphorylation that promotes phosphorylation of intracellular substrates, including the insulin receptor substrate 1/2 (IRS-1/2). When IRS is phosphorylated by the IR, signals are passed down to PI3K, which then phosphorylates the 3' position of phosphatidylinositol-4,5-bisphosphate (PI(4,5)P2) generating phosphatidylinositol-3,4,5-trisphosphate (PI(3,4,5)P3). PIP3 recruits the 3'-phosphatidylinositol dependent kinase 1 (PDK1) to activate protein kinase B (PKB/AKT) which suppresses GSK3 activity by phosphorylating an N terminal serine of the enzyme. The inactivated GSK3 no longer phosphorylates sites 3a, 3b, 3c and 4 at the C-terminus of glycogen synthase and leaves the enzyme active. Since insulin causes dephosphorylation of both the N and C terminal phosphorylation sites [82, 83], it seems that insulin control of GS cannot be explained only by the regulation of GSK-3. It has been proposed that insulin stimulation activates PP1G with R_{GL} [84] or PTG [60, 85]. However, GS is still activated by insulin signaling in R_{GL} knockout mice [86], implicating other targeting subunits of PP1 in insulin stimulated dephosphorylation of GS in muscle. In the diabetic state, impaired muscle glycogen metabolism is associated with deficient insulin action [1].

Epinephrine works antagonistically with insulin in muscle, stimulating glycogenolysis and inhibiting glycogenesis. By interacting with β-adrenergic receptors, epinephrine activates adenylyl cyclase to convert ATP into cAMP. cAMP binds to the regulatory subunit of cAMP dependent protein kinase (PKA) and releases the active catalytic subunit to phosphorylate the α and β subunits of PhK. The activated catalytic subunit of PhK phosphorylates its primary downstream target GPh, shifting it to the more active R state (*a* form), thus promoting glycogenolysis. For glycogen synthesis inhibition, PKA phosphorylates GS at sites 2, 1a and 1b [1], and regulates PhK to phosphorylate site 2. PKA dependent phosphorylation has been proposed to regulate PP1 by phosphorylating R_{GL} leading to dissociation of the catalytic subunit PP1c, thus

reducing the activity of PP1G [87]. A second mechanism for PKA control is by phosphorylating inhibitor-1, a small protein inhibitor of PP1 that is effective only after phosphorylation [88].

1.4.2 Regulation in the liver

In the fed state, glucose follows its concentration gradient into liver via the GLUT2 glucose transporter which not controlled by insulin. Glucokinase and GS play large part in determining the rate of glycogen synthesis [1]. Simultaneously, allosteric binding of glucose to the phosphorylated form of glycogen phosphorylase makes it a better substrate for PP1G-G_L dephosphorylation. This provides a negative feedback control on glycogen degradation which also releases glucose [89]. Moreover, the glucose induced transition of phosphorylase *a* to the *b* form relieves its inhibition on PP1G-G_L, leading to dephosphorylation and activation of glycogen synthase [89].

When glucagon action predominates over insulin action in response to decreased blood glucose, liver first breaks down glycogen reserves before activating gluconeogenesis. Glucagon activates liver adenylyl cyclase to increase cAMP levels which then promotes PKA dependent phosphorylation of GS, leading to its inactivation. At the same time, PhK is activated by PKA phosphorylation resulting in phosphorylation and activation of phosphorylase. In addition, activated phosphorylase (*a* form) inhibits GS activity by binding G_L to modulate PP1G-G_L phosphatase activity towards GS [58, 90]. Glucose-1-P is released by glycogen phosphorolysis and then converted to G6P by phosphoglucomutatse. Glucose-6-phosphatase (G6Pase) converts G6P to glucose that can help maintain blood glucose levels and provide an energy supply for other tissues, such as the brain, where glucose is a critical energy source.

2. Diseases associated with glycogen metabolism

Glycogen accumulation and consumption is closely correlated with the glucose metabolism and usage in individual cells as well as in the whole body. Abnormal glycogen storage can be considered as both a result and a cause of several diseases. Abnormalities in glycogen metabolism are linked to a number of diseases, including single gene defects like glycogen storage diseases, and more complex, multigenic disorders like type 2 diabetes.

2.1 Glycogen storage disease type 0: Glycogen synthase deficiency

Glycogen storage disease type 0 (GSD-0) was first identified as a GS deficiency resulting in infantile fasting hypoglycemia [91] coincident with high blood ketones, low alanine and lactate concentrations, as well as postprandial hyperglycemia, hyperlactatemia and hyperlipidemia [92]. GSD-0 is caused by autosomal recessive mutations in the *GYS2* gene, which encodes the liver isoform of glycogen synthase. In *Gys2* null mice, liver glycogen is reduced by 95%. These mice have mildly hypoglycemia, elevated basal gluconeogenesis and impaired insulin suppression of endogenous glucose production [93]. Mutation in the *GYS1* gene, which encodes the muscle isoform of GS, was reported to cause the sudden cardiac death of an eight year old patient who collapsed during a bout of exercise [94]. Consistent with this observation, about 90% of *Gys1* knockout mice died soon after birth most likely due to cardiac dysfunction [95].

2.2 Glycogen storage disease type I: von Gierke's disease; Glucose-6-phosphatase deficiency; Hepatorenal glycogenosis

Glycogen storage disease type I (GSD-I) is autosomal recessive and has four subtypes corresponding to different defects in the G6Pase system. GSD-Ia is caused by the defect of G6Pase catalytic subunit (G6Pase-α), encoded by *G6PC*, leading to over-accumulation of glycogen and lipids in liver, kidney and intestinal mucosa. GSD-Ib is caused by the deficiency of G6P transporter (G6P translocase, G6PT). Both GSD-Ia and Ib patients have hypoglycemic seizures, hepatomegaly, growth retardation and life-threatening lactic acidosis [96, 97]. In gluconeogenesis and glycogenolysis, G6P is translocated from the cytoplasm into the lumen of the ER by G6PT, where G6Pase hydrolyses it [97]. The released glucose is transferred into blood stream. The failure of G6P dephosphorylation contributes to elevated levels of intracellular G6P that drives excessive glycogen accumulation accompanied by severe hypoglycemia [98]. Both GSD-Ic and -Id are deficient in liver microsomal transport system, in which GSD-Ic lacks phosphate mobilization while GSD-Id is unable to transport glucose [99]. Most GSD-Ic and all GSD-Id patients reported so far have mutations on the same gene encoding G6PT as in the GSD-Ib patients, but are diagnosed differently in clinical and biochemical views. Deficiency of a new microsomal phosphate transporter, the Na(+)/phosphate co-transporters 4 (NPT) has been suggested to affect some of the GSD-Ic patients [100].

2.3 Glycogen storage disease type II: Pompe disease; Acid α-glucosidase deficiency; Acid maltase deficiency; α-1, 4-glucosidase deficiency

A lysosomal storage disease, glycogen storage disease type II (GSD-II) is an autosomal recessive disorder caused by a deficiency in the lysosomal GAA. Based on the severity and onset time, patients of GSD-II can be divided into infantile, non-typical infantile, juvenile and adult forms [101]. In the infantile GSD-II, GAA is absent in all tissues, with severe hypotonia, cardiomegaly,

macroglossia, and mild hepatomegaly. Glycogen over accumulation in the lysosomes of the cardiac and skeletal muscle, usually causes death within a year after birth. The non-typical infantile type patients can survive longer than 12 months with less severe cardiomyopathy and ventricular hypertrophy [102]. The juvenile form starts in the first decade, is associated with residual GAA activity and is accompanied by weakness in skeletal and respiratory muscles, and mild hepatomegaly [103]. The adult phenotype, usually defined by the onset in the third to sixth decade of life, is similar to the juvenile form but with higher levels of residual GAA activity and a slower development of skeletal muscle weakness [104].

The phenotype of the *Gaa* null mice approximately reproduces the human disease but with later progression of symptoms comparing to the onset of the human disease [105]. Crossing *Gaa* knockout mice with a transgenic mouse overexpressing a hyperactivated form of GS sped up the disease progression and generated structurally abnormal polyglucosan, similar to what is seen in patients with Lafora disease, glycogen storage disease type IV (Andersen's disease) and type VII (Tarui's disease) [66]. A mouse *Gaa* and *Gys1* double knockout showed significant decrease in glycogen levels of the cardiac and skeletal muscle, a notable decline in lysosomal enlargement and formation of autophagic vesicles as well as a complete correction of cardiomegaly and improved exercise capacity [106]. These data suggest long-term inhibition of GS as a new approach to GSD-II treatment. Treatments for GSD-II include recombinant GAA infusion which led to the approval of enzyme replacement therapy in 2006, and development of experimental therapies, such as adeno-associated virus vector-mediated gene therapy [99].

2.4 Glycogen storage disease type III: Cori disease; Forbes disease; Amylo-1,6-glucosidase deficiency; Glycogen debrancher deficiency

Glycogen storage disease type III (GSD-III) is an autosomal recessive disorder caused by mutations in the *AGL/DBE* gene leading to defects in the glycogen DBE (AGL) [107]. GSD-IIIa affects both the liver and muscle in 80% of the GSD-III cases and GSD-IIIb affects only the liver in nearly 15% [108, 109]. In rare cases, GSD-IIIc and GSD-IIId have also been reported as selective deficiencies in glucosidase and transferase activities respectively [110, 111]. Symptoms usually present during infancy with hepatomegaly, hypoglycemia, short stature, dyslipidemia, and in a few cases, slight mental retardation. The aberrant glycogen has very short outer chains because the complete phosphorolysis by GPh requires the elimination of branch points.

2.5 Glycogen storage disease type IV: Andersen's disease; Brancher deficiency; Amylopectinosis; Glycogen branching enzyme deficiency

Mutations in the *GBE1 gene* which encodes branching enzyme cause glycogen storage disease type IV (GSD-IV) which is autosomal recessive and can be one of the most severe of the glycogen storage disorders [112]. The phenotypes of GSD-IV are very heterogenous, partly because of the existence of tissue specific isoenzymes. In the canonical hepatic form, patients present with hepatosplenomegaly, progressive liver cirrhosis and failure to thrive [113]. In the multiple system involvement, defects in both muscle and liver are detected, including peripheral myopathy with or without cardiomyopathy, neuropathy, and liver cirrhosis [114]. The severity of the disease correlates with the polyglucosan accumulation level [115]. The polyglucosan has long unbranched chains resembling amylopectin structure, accompanied by a loss of solubility, eventually driving cellular dysfunction and degeneration. Liver transplantation is the only effective treatment so far for GSD-IV patients with progressive liver disease [99].

2.6 Glycogen storage disease type V: McArdle's disease; Myophosphorylase deficiency; Muscle glycogen phosphorylase deficiency

Glycogen storage disease type V (GSD-V) is caused by mutations in the *PYGM* gene encoding the muscle form of GPh. Symptoms usually start in young adulthood, mild and with exercise intolerance and muscle cramps due to the lack of glucose release from glycogen degradation as fuel [116].

2.7 Glycogen storage disease type VI: Hers disease; Liver glycogen phosphorylase deficiency

Glycogen storage disease type VI (GSD-VI) is a rare form of GSD, caused by a deficiency in the the liver GPh which is encoded by *PYGL* gene. Patients manifest infantile hepatomegaly and growth retardation with symptoms usually relieved as the patients age [99].

2.8 Glycogen storage disease type VII: Tarui disease; Muscle phosphofructokinasedeficiency; Glycogen storage disease of muscle

The symptoms of glycogen storage disease type VII (GSD-VII) are muscle cramps and myoglobinuria with exercise similar to that observed in GSD-V (McArdle's disease). GSD-VII is autosomal recessive and caused by mutations in the *PFKM* gene, encoding the muscle isoform of phosphofructokinase. Phosphofructokinase (PFK) (EC 2.7.1.1.11) phosphorylates fructose-6-phosphate to generate fructose-1,6-bisphosphate. This reaction is a key regulatory step in glycolysis. The deficiency in PFK causes accumulation of glycolytic intermediates that builds up G6P, the allosteric activator of GS, thus resulting in excessive glycogen synthesis [117].

2.9 Phosphorylase activation system defects: Phosphorylase kinase system defects; GSD-VIa, IX, X, or VIII)

PhK contains four subunits (α, β, γ and δ) that are encoded by different genes on different chromosomes with tissue specific expression. The α and β subunits are regulatory, the γ subunit has the catalytic activity, and δ is responsible for Ca^{2+} binding. The two isoforms of the α subunit, muscle and liver, are both encoded by genes located on the X chromosome [118] while the genes encoding other subunits (β, γ and δ) are autosomal. Based on tissue specificity and incidence, there are four main clinical variants, affecting liver only, liver and muscle, muscle only, and heart only [119]. The liver specific cases are categorized into two types, the autosomal recessive form and the X-linked deficiency which is subsequently divided into two subtypes, XLG-I and II. As one of the mildest of GSD, X-linked liver phosphorylase kinase deficiency (XLG; X-linked liver glycogenosis type I and type II, formerly GSD-VIII or GSD-VIa) is similar to GSD-VI since both disorders demonstrate low phosphorylase activity in the absence of adenosine monophosphate (AMP) [99]. The autosomal liver and muscle phosphorylase kinase β-subunit deficiency leads to remarkable accumulation of glycogen in both liver and muscle, and is characterized by a distended abdomen due to significant hepatomegaly [120]. The γ-subunit has two isoforms, the muscle form and the testis/liver form. Mutations in the gene for the liver isoform are linked to higher risk of cirrhosis [121].

2.10 Glycogen storage disease type XI: GSD-XI; Fanconi-Bickel Syndrome, FBS

Glycogen storage disease type XI (GSD-XI) is a rare, autosomal recessive transmitted disorder, caused by defects in a glucose transporter, GLUT2. The disease is defined by hepatorenal glycogen accumulation, fasting hypoglycemia accompanied by postprandial hyperglycemia and hypergalactosemia, proximal renal tubular dysfunction, rickets and growth retardation [99].

2.11 Lafora disease

Lafora disease (LD) is an autosomal recessive progressive myoclonic epilepsy, rare but lethal, with the onset usually in the second decade of life [122]. It is a neurodegenerative disorder and can also be considered a glycogen storage disease due to the formation of polyglucosan inclusion bodies with insoluble, poorly branched glycogen-like structures, known as Lafora bodies, in neural and other tissues. The condition is characterized by epilepsy, myoclonus and dementia. Mutations in either the *EPM2A* gene encoding a dual specificity phosphatase laforin or *NHLRC1* (also known as *EPM2B*) encoding an E3 ubiquitin ligase malin cause ~90% cases of LD [123]. The remaining cases may be caused by mutations in a third, unknown gene, or in the regulatory elements of the two known genes [124, 125]. Therapeutic approaches for LD at present are limited to symptomatic management of the epilepsy, myoclonus and other complications. Most recent research on LD has been aimed at understanding the functions and relationships of laforin and malin in the disease progression, with the hope of finding future therapies [126].

3. Carbohydrate-binding modules

3.1 Carbohydrate-binding module families

A carbohydrate-binding module (CBM) is described as a contiguous amino acid sequence folding as a separate domain and with non-catalytic carbohydrate-binding activity [127]. CBMs are usually present within large carbohydrate-active enzymes, which separate this class of carbohydrate-binding protein from other non-catalytic sugar binding proteins like lectins and sugar transport proteins. However, there are a few exceptions, such as CBMs in cellulosomal scaffolding proteins and examples of independent putative CBMs [128].

CBMs were first discovered as modules that bound cellulose from the bacterium *Cellulomonas fimi* [129] and the fungus *Trichoderma reesei* [130]. Thus, they were previously classified as cellulose-binding domains (CBDs). However, since non-catalytic modules derived from glycoside hydrolase (GH) families were identified in 1999 that bound carbohydrates other than cellulose, the more inclusive term CBM has been used to reclassify these polypeptides [131]. CBMs generally come with carbohydrate hydrolases degrading insoluble polysaccharides. Several CBM families include proteins that bind storage polysaccharide like starch and glycogen [131].

Since more and more modules are continually being discovered, the CBMs have been classified into different families based on amino acid sequence similarity, binding specificity and structure [127, 131]. Comprehensive classification data of CBMs are collected in the carbohydrate active enzymes database (CAZy) [132] (http://www.cazy.org/Carbohydrate-Binding-Modules.html).

Up to 2010, 61 families of CBMs have been reported. Among them, three-dimensional structures of representative members of 55 families have been determined. The β-sheet turns out to be the most prevalent conformational element in the architecture of CBMs. Based on the structures, CBMs families have been grouped into 7 fold families: β-sandwich, β-trefoil, oligonucleotide /oligosaccharide binding (OB) folds, hevein fold, cystein knot, unique and unique with hevein-like fold [128, 131, 133] (Table 1). (1) The β-sandwich is the dominant fold group, which contains two β-sheets comprising 3-6 antiparallel β-strands [134]. Based on their fold, this group is further divided into two sub-families, the β-jelly roll fold and the immunoglobulin fold [128]. (2) The β-trefoil, which appears second in frequency, forming 6 hairpin turns by 12 strands of β-sheet. The overall fold, having a pseudo-3-fold symmetry, consists of a 6-stranded β barrel with a triangular hairpin triplet cap [135]. (3) The OB fold shows a closed beta-barrel formed by a coiled 5-stranded β-sheet. This barrel is capped by an α-helix located between the third and fourth strands [136]. (4) The

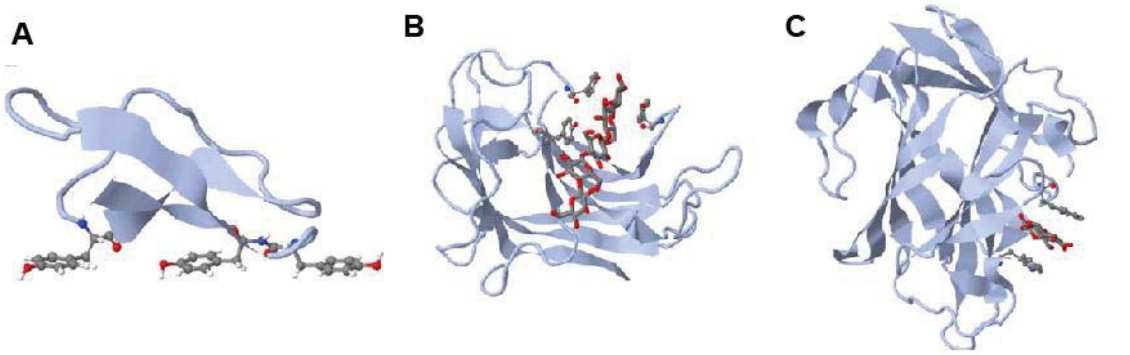
hevein fold contains primarily coils along with 2 small β-sheets and a small helix region[131].

Fold	**CBM families**
β-sandwich	2, 3, 4, 6, 9, 11, 15, 16, 17, 20, 21,22, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 40, 41, 42, 44,47,48, 51
β-trefoil	13, 42
cystein knot	1
Unique	5,12
OB fold	10
Hevein fold	18
Unique (hevein-like fold)	14

Table 1. CBMs classification based on structure (fold).

Adapted from Guillen D., Sanchez S. and Rodriguez-Sanoja R. Appl Microbiol Biotechnol (2010) [133].

CBMs have also been classified into three types by functional similarity in ligand-binding [128, 131, 133]. (1) Type A includes “surface-binding” CBMs, which bind to insoluble, highly crystalline polysaccharides like cellulose and/or chitin. The aromatic amino acid residues at the binding sites have a planar conformation that interacts with the flat surface of cellulose or chitin crystals. There is little or no affinity of type A class CBMs to soluble carbohydrates. (2) Type B are “glycan-chain-binding” CBMs. Their binding sites are described as grooves or clefts comprising several subsites where the aromatic residues interact with individual polysaccharide chains. The aromatic side chains form twisted or sandwich platforms for ligand binding. The binding affinity is determined by the degree of polymerization of the carbohydrate ligand. Type B CBMs are equipped to prefer single glycan chains over crystalline surfaces for interaction. (3) Type C are “small sugar binding” CBMs. They are lectin-like and bind to mono-, di-, or trisaccharides due to steric restriction in binding sites. (Figure 3)



Adapted from Guillen D, Sanchez S and Rodriguez-Sanoja R.
Appl Microbiol Biotechnol. (2010)

Figure 3. Structures of CBMs classified into three types by functional similarity [133].

(A) Type ACBM from *Trichoderma reesei* cellobiohydrolase I. (B) Type B CBM from *Cellulomonas fimi* endo-1,4-glucanase C. (C) Type C CBM from *Thermotoga maritime* xylanase 10A.

Studies of residues involved in binding sites suggest that aromatic amino acids, typically tryptophan, but also highly conserved tyrosine residues play a critical role in ligand binding. And the ligand specificity is determined by the orientation of aromatic residues. The binding structure is stabilized by strong Van der Waals force accumulated from interactions between aromatic residues and sugar rings. Hydrogen bonds formed by the side chain of other polar amino acid residues may also help the stabilization. Proper topology of binding sites is responsible for efficient polysaccharide ligand recognition and binding [137]. Therefore, subtle changes lead to different ligand specificity in CBMs even with apparent similar structures.

In a given protein, the localization of CBMs can either be C-terminal or N-terminal, and occasionally central, in between catalytic modules. CBMs contain from 30 to about 200 amino acids, and exist as a single domain, or double or triple units arranged in tandem [127, 133]. CBMs are also discovered as single proteins by themselves, for instance, Ole e10 from olive pollen [138], CBP21 from *Serratia marcesens* [139], and E7 and E8 from *Thermobifida fusca* [140].

Since carbohydrates are involved in various biological processes such as metabolism, energy storage, structural support, antibiosis, immunological functions etc, it is very important to understand the mechanism of protein-carbohydrate interactions. The key role of CBMs is carbohydrate recognition and specific binding. In hydrolytic proteins, CBMs contribute to increase catalytic rates by introducing the catalytic module into the proximity of substrate and prolong the enzyme association with substrate. In nonhydrolytic proteins, CBMs can serve as part of a scaffolding unit as in cellulosome. Some CBMs have been reported to be capable of nonhydrolytically disrupting polysaccharide structures [127, 131]. Studies also indicated that CBMs worked on surface/interfacial modification of cellulose fibers [127]. So far, the most common biological function of CBMs is to bind polysaccharide and promote efficient hydrolysis. However, other CBM-related functions have been continually discovered, such

as development and modification of plant cell wall, pathogen defense, virulence, and etc. As more structures and families of CBMs have been discovered and studied, our knowledge in CBMs expands and practical applications of CBMs have been reported in different biotechnology fields, such as bioprocessing, targeting, cell immobilization, CBM engineering, analytical tools in research and diagnostics, bioremediation, and fiber modification [127].

3.2 Starch-binding domain

Starch-binding domain (SBD), a group of CBMs that bind starch, has nine families at present, including CBM20, CBM21, CBM25, CBM26, CBM34, CBM41, CBM45 [141, 142], CBM48 and CBM53 [143]. Three-dimensional structures of all but CBM45 and CBM53 are known. SBDs have been classified into different CBM families based on the diversity of amino acid sequences. Despite the low sequence similarities, all SBDs analyzed display a similar architecture described as β-sandwich fold with immunoglobulin-like topology [143].

3.2.1 CBM20

The earliest discovered SBDs were in the CBM20 family, which has been the most studied family in this clan so far. SBD was considered exclusively associated with microbial amylolytic enzymes for many years until more extensive sequence databases became available. Sequences similar to the CBM20 family of SBD were identified by BLAST searches. Orthologs or putative SBD sequences have been identified in a broad range of bacterial, fungal, archaea, plant or animal proteins. There are nearly 400 records in the CAZy database of CBM20s, in which seven are from archaea, 231 are from bacteria, 138 are from eukaryote and 21 are unclassifed. These proteins can be intracellular or secreted, amylolytic or nonamyloytic. Although these CBM20s belong to proteins with different catalytic modules and from diverse phylogenic origins, they are generally involved in starch or glycogen metabolism. Usually

containing 90-130 amino acids, CBM20s have no absolutely conserved residues in the family by comprehensive sequence studies [141, 142, 144].

The three-dimensional structures of CBM20s display a twisted and open-sided β-barrel structure built by several β-strands [145, 146] (Figure 4A). Two individual starch-binding (or glucan-binding) sites are defined [142, 147]. The structure of the best studied CBM20, the glucoamylase (GA) SBD from the fungus *Aspergillus niger*, was solved by nuclear magnetic resonance (NMR) in a free state or bound to β-cyclodextrin [146, 148]. Binding site 1, small and shallow with inflexible structure on binding, comprises W543, K578, W590, E591 and N595. The aromatic rings of W543 and W590 form a rigid and shallow hydrophobic platform for stacking two units of α-1, 4- glucoses. There is little conformational change of site 1 in the ligand bound state. Binding site 2 consists of T526, Y527, G528, E529, N530, D554, Y556 and W563, covering a broader surface area for ligand binding and requiring significant conformational change. The pivotal aromatic residues Y527 and Y556 stick out from protein surface, quite different from site 1 where aromatic rings lie parallel. Although W563 is unable to interact with ligand due to its lack of surface accessibility, it is critical in keeping contacts with different residues in and around binding site 2. Since the two binding sites differ in structure and flexibility, a different functional role is suggested for each the site. While Site 1 is indicated as the primary recognition site for starch, site 2 is suggested as the specific site to lock SBD on the starch surface. Studies in cyclodextrin glucanotransferase (CGTase) [149], α-amylase [150] and β-amylase [151] also provided supportive data.

Based on the tree-dimensional structures, consensus amino acids associated with starch binding have been analyzed. At the beginning, the sequence alignment of eight species from bacteria and fungi revealed four areas of sequence similarity [152]. Eleven invariant amino acids were identified including three tryptophans, three glycines and single residues of proline, threonine, lysine, asparagine and leucine. The conserved tryptophans were considered involved in

association to glucosyl residues via Van der Waals force and/or hydrogen bonding [137]. The glycines and proline, as well as threonine, lysine and asparagines were suggested to maintain the conformation of the polypeptide chain [152]. Nevertheless, not all those conserved amino acids appeared to have same importance with expanded research in broader collections of CBM20s. Binding site 1 carries two conserved trptophans, e.g. W543 and W590 in GA SBD from *A. niger* [153], as well as W643 and W689 (or W616 and W662 after the deletion of 27 amino acids of signal peptide) in CGTase SBD from the bacteria *Bacillus circulans* strain 251[154]. The most conserved region in the CBM20 motif is around the former tryptophan including residues LGxW. These two tryptophans are not entirely invariant, but in the 103 sequences analyzed, none of them lacks both residues [142]. Binding site 2 contains one tryptophan aligning with W563 in GA SBD from *A niger*, and two tryrosines in some cases. One tyrosine aligning with Y527 in GA SBD from *A. niger* is found in 23.3% of the sequences analyzed, and the other one aligning withY556 shows in 43.7% of the sequences [143]. Besides the earlier mentioned 11 consensus residues, one phenylalanine aligning with F519 in GA SBD from *A niger* is conserved in 87.5% of the analyzed sequences, including SBDs from outside of amylolytic enzymes, e.g. animal laforin and starch-binding domain-containing protein1 (Stbd1). Different residues are found at this position such as leucine, isoleucine or valine in several bacterial β-amylase, and arginine in glucan, water dikinase 3 (GWD3) from the plant *Arabidopsis thaliana* [143]. Another highly invariant tryptophan, aligning with W615 in GA SBD from *A niger* is considered primarily for conformational maintenance, because the W615K mutant had poor stability [153].

CBM20s were first discovered as C-terminally positioned. However, some plant GWD3 and 4-α-glucanotransferases, in addition to a number of unknown eukaryotic proteins in sequence alignment studies, have an N-terminal CBM20. Therefore, CBM20s are described as prominently but not exclusively localizing at C-termini [143].

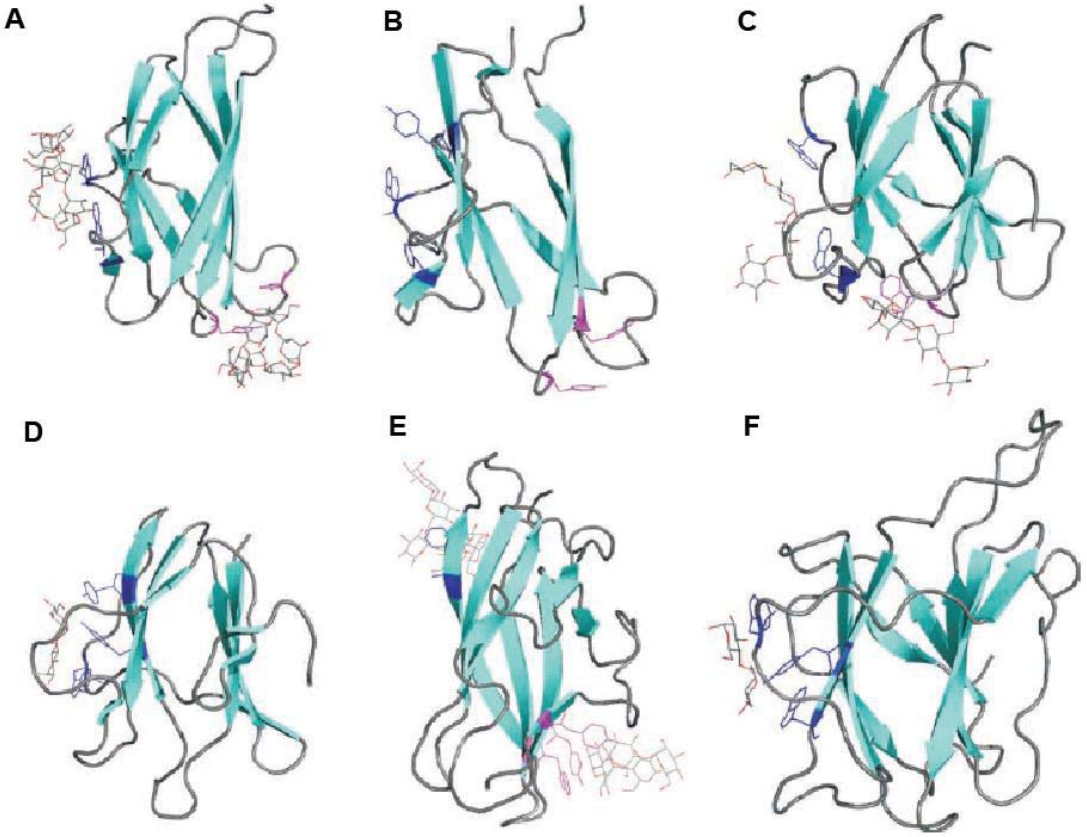
3.2.2 CBM21, CBM48 and CBM 53

CBM21s are mainly N-terminally positioned and 90-140 amino acids long. The representative CBM21 is in GA from *Rhizopus oryzae* [152] and one single lysine residue, K34 is invariant through the CBM21 family alignment [142]. The solved CBM21 structure demonstrates a β-sandwich fold and immunoglobulin-like composition [155] (Figure 4B), similar to the canonical CBM20 architecture. CBM21 has not been studied as intensely as CBM20, partly because the situation is more complicated in CBM21. Most members in the CBM21 family are from unknown proteins of miscellaneous origins. Only 10% of the sequences in this family are from microbial amylolyctic enzymes while a considerable number of the rest are eukaryotic protein phosphatases and/or their regulatory subunits, e.g. PTG involved in glycogen metabolism [51]. CBM21 has been suggested to form a clan with CBM20 through bioinformatic analysis [142]. Despite infrequent slight variations in structure, close evolutionary relationships among CBM20, 21, 48 and 53 has been proposed by the alignment of 60 selected SBD sequences. Besides SBDs linked with catalytic subunits of bacterial or fungal GH families, protein domains in higher eukaryotes have also been analyzed. The mammalian regulatory subunit (β-subunit) of AMPK [156] and the plant starch excess 4 protein which has been reported to be a laforin-like phosphoglucan phosphatase [157, 158], are members of CBM48, and the three SBD repeats in the plant starch synthase III from *A. thaliana* have been included in the new CBM53 family.

3.2.3 CBM25, CBM26, CBM34, CBM41 and CBM45

CBM25 and CBM26 are structurally related, and both exist in the maltohexaose-forming amylasefrom *Bacillus halodurans*. The SBDs in CBM25 can be single or double units, while the CBM26 members usually appear as tandem repeats. CBM25 has two binding sites while CBM26 bears only one [141] (Figure 4C and D). CBM34 family covers the N-terminal modules of the neopullulanase subfamily like neopullulanase, maltogenic amylase and cyclomaltodextrinase,

which are all closely related members of the GH13 family (Figure 4E). CBM41 contains mostly bacterial pullulanases of GH13 family as well as many undefined proteins [141] (Figure 4F). A number of those pullulanases are from human pathogens with structures similar to that of CBM20 [159, 160]. On the contrary, members in CBM45 family are mainly originated from eukaryotic proteins, as the tandem N-terminal domains of plant α-amylases and α-GWD [141].



Adapted from Machovic M and Janecek S. Cell Mol Life Sci. (2006)

Figure 4. The structural features of SBD from individual CBM families [141]. (A) The CBM20 from *Aspergillus niger* glucoamylase in complex with β-cyclodextrin. (B) The CBM21 from *Rhizopus oryzae* glucoamylase. (C) The CBM25 from *Bacillus halodurans* maltohexaose-forming amylase. (D) The CBM26 of *Bacillus halodurans* maltohexaose-forming amylase. (E) The CBM 34 of *Thermoactinomyces vulgaris* 'α-amylase' TVA I. (F) The CBM41 of *Klebsiella pneumoniae* pullulanase. The tryptophan and tyrosine residues involved in substrate interaction are colored blue in the binding site 1 and magenta in the binding site 2.

4. Starch-binding domain-containing protein 1

Starch-binding domain-containing protein 1 (Stbd1), sometimes called genethonin 1, is a novel protein with unknown function. The chromosomal location of the human gene is 4q24-q25. Stbd1 was first identified by Bouju *et al.* in 1998 [161], as a striated muscle-enriched protein encoded by a novel gene registered in the Genexpress index. The cDNA clone was termed GENX-3414, and was isolated from a human skeletal muscle cDNA library with a set of 1091 cDNA clones [162]. Bouju *et al.* selected a 2117-bp cDNA clone from the library as a northern blot probe and an mRNA about 2.4 kb was detected in different human tissues with high level in skeletal and cardiac muscles, and moderate levels in liver and placenta. No hybridization with mRNA was detected in pancreas, kidney, lung and brain according to the report. With successful isolation and sequencing of the full-length cDNA for the human GENX-3414 transcript, their studies of the protein product provided some initial information. From sequence analysis, they pointed out a hydrophobic segment at the N-terminus, a possible internal leucine-zipper motif, and several potential post-translational modification sites, such as N-linked glycosylation, myristoylation and phosphorylation. Immunoblot of human tissue extracts from skeletal muscle, heart and placenta showed a single or double-band protein with the apparent molecular masses of 41-43 KDa. The protein was detected at high concentrations in membrane-enriched preparations of muscle, but not in the purified myofibrillar preparations. Immunofluorescence analysis showed the protein was present within the cytoplasm and not at the peripheral sarcolemmal membrane. These data so far suggested the protein was preferentially localized to the internal membrane systems (T-tubules and junctional sarcoplasmic reticulum) of the skeletal muscle.

In 2002, Janecek [163] described the same protein which he referred to as genethonin, in a bioinformatics analysis focusing on the conserved starch-binding domain (SBD). The growing databases raised the interest of searching SBD-

related sequences in proteins other than amylase and of non-microbial origin, especially when an SBD-like sequence was reported at the N-terminus of laforin [164], a human protein involved in the progressive myoclonic epilepsy, Lafora disease. By BLAST searches, human Stbd1 (genethonin) was identified to contain higher than 25% identity to SBD from the CBM20 family and has the classical C-terminal position of the SBD found in microbial amylases [163]. The tryptophan residues that are important to the structure and function of SBD from *B. circulans* strain 251 CGTase (W616, W 636, W662 and W684) [145, 148, 154] and *A. niger* GA (W563, W590, W615) [153, 165] are conserved in Stbd1(W293, W307, W334, W355) [163]. Homology modeling of 3-dimentional structure features of the C-terminus of human Stbd1 showed strong resemblance to the template structure of SBD from Bacillus sp. Strain 1011 CGTase (Figure 5), in the number and mutual orientation of β-strands and the side-chain orientation of the three tryptophans involved in starch-binding sites 1 and 2 [163].

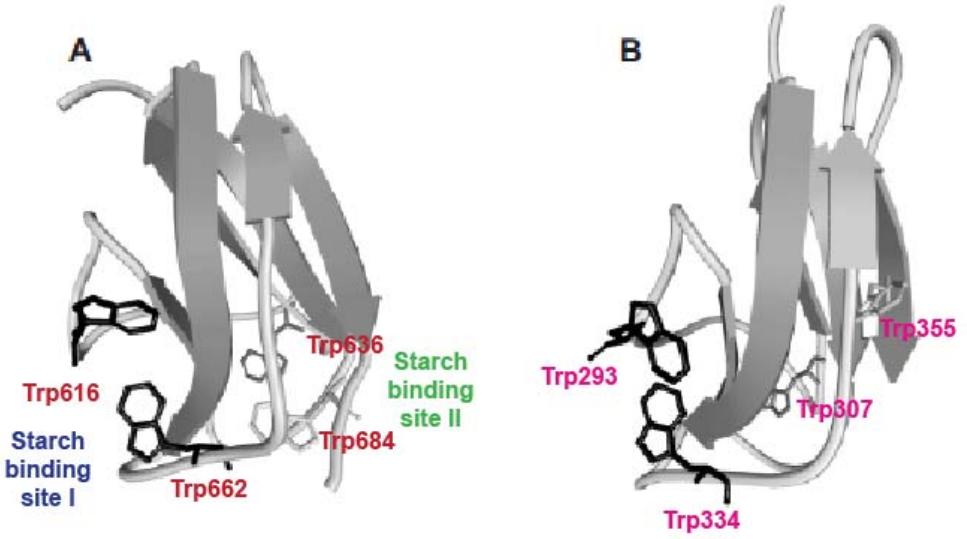


Figure 5. Structures of conserved SBDs [163].

(A) X-ray crystal structure of the Starch-binding domain from *Bacillus.* sp strain 1011 cyclodextrin glucanotransferase (CGTase). Starch-binding site 1 includes W616 and W662; Starch-binding site 2 includes W636; W684 serves primary structural role. (B) Homology modeling structure of the C-terminal part of human Stbd1.

Although a carbohydrate-binding function of Stbd1 was indicated by the bioinformatics analysis, little work has addressed this protein until recently. In 2010, Stbd1 was identified by two different proteomic analyses. Stapleton *et al*. [166] analyzed hepatic glycogen particle and identified Stbd1 as a protein associated with glycogen in both mouse and rat liver. Behrends *et al*. reported Stbd1 as one of 67 high-confidence candidate interaction proteins (HCIPs) in an autophagy interaction network in human cells [167]. Our data on Stbd1 structure, subcellular distribution and functional analysis (see Results) provide more detailed insight into this novel protein [168].

5. Autophagy

Autophagy, or autophagocytosis, is a collection of catabolic processes involving the delivery and degradation of a portion of the cytoplasm through the lysosomal machinery (Figure 4). The process involving bulk segregation of cellular constituents in mammalian cell was first reported in 1957 [169]. Ultrastructural changes were observed in rat liver following glucagon injection, and based on that, autophagy was first introduced as a term by De Duve *et al*. in the 1960s [170, 171].

Although autophagy has been studied in mammalian cells for many years, insight into the molecular mechanism was first revealed in the yeast *Saccharomyces cerevisiae* which is genetically tractable. The yeast vacuole, comparable to the lysosome in mammalian cells, is an acidic compartment containing hydrolytic enzymes. Since 1997, 32 autophagy-related (ATG) genes have been identified in yeast, named from *ATG1* to *ATG32*. Most *ATG* genes are involved in autophagosome formation and are well conserved from yeast to mammalian systems [172]. The proteins encoded by *ATG* genes (Autophagy-related gene) are shared in nonspecific autophagy, specific and biosynthetic cytoplasm-to-vacuole targeting pathway (Cvt), and selective breakdown of peroxisomes and mitochondria [173].

A variety of autophagic processes exist, all having in common the degradation of intracellular components via the lysosome. According to the ways substrates reach the lysosome, three major forms of autophagy have been described: macroautophagy, microautophagy and chaperone-mediated autophagy [174] (Figure 6).

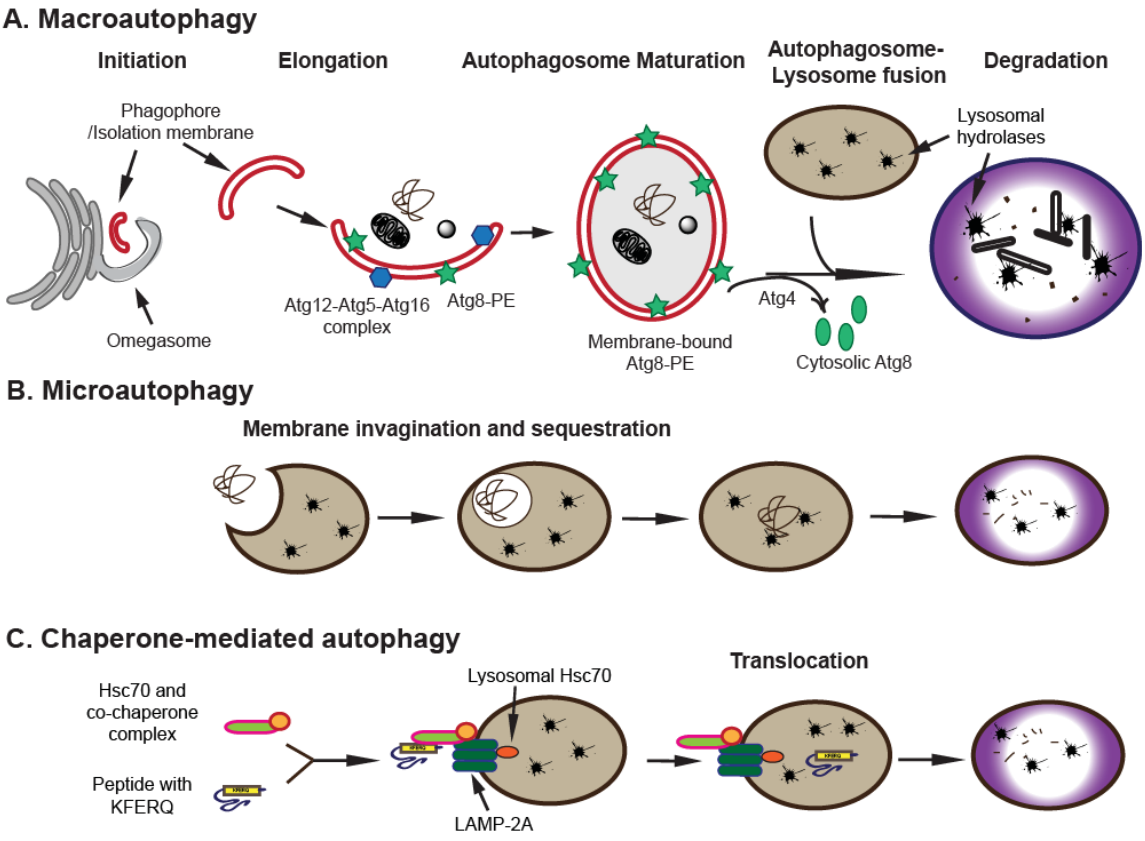


Figure 6. Schematic models of three forms of autophagy.

(A) Macroautophagy. (B) Microautophagy. (C) Chaperone-mediated autophagy.

5.1 Macroautophagy

Macroautophagy, usually just referred to as autophagy, involves the formation of an autophagosome, a double -membrane structure that surrounds a targeted region of the cell, separating the contents from the rest of the cytoplasm. The resultant vesicle then fuses with a lysosome that subsequently degrades the contents. ER is generally considered as one source of the sequestering membrane, whereas the phagophore or preautophagosome (also called isolation membrane), a unique hypothetical compartment is thought as membrane donor, [175, 176]. The macroautophagy process can be divided into several steps: omegasome (shaped like Ω) formation, the initiation and elongation of isolation membrane/phagophore, autophagosome formation, autophagosome-lysosome fusion, and degradation [172] (Figure 6A).

5.1.1 Omegasome formation

mTOR, a serine/threonine kinase, negatively regulates autophagy by phosphorylating Atg13, a component of ULK protein kinase complex which is required for autophagy induction and activity. The complex contains unc-51 kinase like protein (ULK, serine/threonine kinase, Atg1), Atg13, focal adhesion kinase family interacting protein of 200 KDa (FIP200) and a new mammalian autophagy protein Atg101. The omegasome, shaped like the Greek letter Ω, is from the ER and contains the double FYVE-domain (Phe-Tyr-Val-Glu) containing protein 1 (DFCP1) which resides in ER and Golgi under nutrient-rich conditions. Binding to phosphatidylinositol-3-phosphate (PI(3)P), the DFPC1-positive omegasome formation is regulated by a class III PI3K complex, which includes Atg14, Vps34 (the catalytic subunit of PI3K), Vps15 (the regulatory subunit of PI3K) and beclin 1 (Atg6) [172].

5.1.2 Initiation and elongation of isolation membrane

The isolation membrane is initiated inside the ring of the omegasome and its elongation requires the protein complex composed of the transmembrane protein Atg9 and the hypothetical WD-repeat protein interacting with phosphoinositides (WIPI-1, Atg18) [172].

5.1.3 Autophagosome formation

The autophagic machinery is conserved in the higher eukaryotes, therefore, breakthroughs in the study of yeast systems provide valuable insights into mammalian systems [177]. It has been reported that the human orthologs of Atg12 and Atg5 showed conjugation in a similar system to that of the yeast (see below for details described in Atg12 conjugation systems) [178].
The formation of autophagosomes involves three steps: nucleation, expansion and completion [179]. The central machinery of autophagosome formation includes two ubiquitin-like (UBL) conjugation systems involving two UBL proteins, Atg12 and Atg8 (which is microtubule-associated protein 1 light chain 3, MAP1LC3 or LC3, LC3B in mammalian systems) respectively [177, 179]. The Atg12-Atg5-Atg16 complex is recruited to the isolation membrane mainly at the external surface and brings Atg8/LC3 to the site where phosphatidylethanolamine (PE)-conjugated Atg8/LC3 helps the expansion of the isolation membrane initiation. When autophagosome formation is completed, the Atg8/LC3 is cleaved off from the PE containing outer surface by the protease Atg4, and Atg12-Atg5-Atg16 complex also dissociates from the completed vesicle [179].

In the Atg12 conjugation system, the hydrophilic Atg12 forms a covalent complex with Atg5 through G186, the C-terminal amino acid of Atg12 and K149 in the center of Atg5 [180]. Atg12 is first activated by Atg7, an E1-like enzyme. A thioester bond is formed between the C-terminal G186 of Atg12 to C507 of Atg7.

Then, Atg12 is transferred to C133 of an E2-like enzyme, Atg10, also forming a thioester bond [181]. Atg12 is ultimately conjugated to Atg5 via an isopeptide bond at K149. Since no enzyme is recognized to cleave that bond, this conjugation is considered to be irreversible. No E3-like enzyme is involved in this process. A small protein Atg16 further binds to Atg5, and then Atg12-Atg5-Atg16 forms a multimeric complex, possibly a tetramer, by homo-oligomerization via the coil-coil region in Atg16 [182]. (Figure 7)

In the Atg8 conjugation system, the 117 residue basic protein Atg8 is modified at its C-terminal residue, by conjugation with PE [183]. First, R117 of the nascent protein is cleaved by a cysteine protease Atg4. Then a thioester bond is formed between the exposed G116 on Atg8 to the C507 on Atg7, the E1-like enzyme. Next, Atg8 is delivered to Atg3, the E2-like enzyme, still through a thioester bond at C234. Finally, PE is attached to G116 of Atg8 by an amide bond and is targeted to membranes. It has been reported that the Atg12-Atg5 conjugate is required for Atg8 to localize to phagophore-assembly sites (pre-autophagosome structure, PAS) [184] and for Atg8-PE conjugation, possibly by an E3-like activity [185]. But no HECT or RING finger domain, which is typical of E3 ligases, has been identified in Atg12-Atg5 [179]. The conversion of free Atg8 to Atg8-PE proceeds when autophagy is induced. The lipidation is reversible as Atg8-PE is able to release free Atg8 by Atg4 cleavage [186]. (Figure 7)

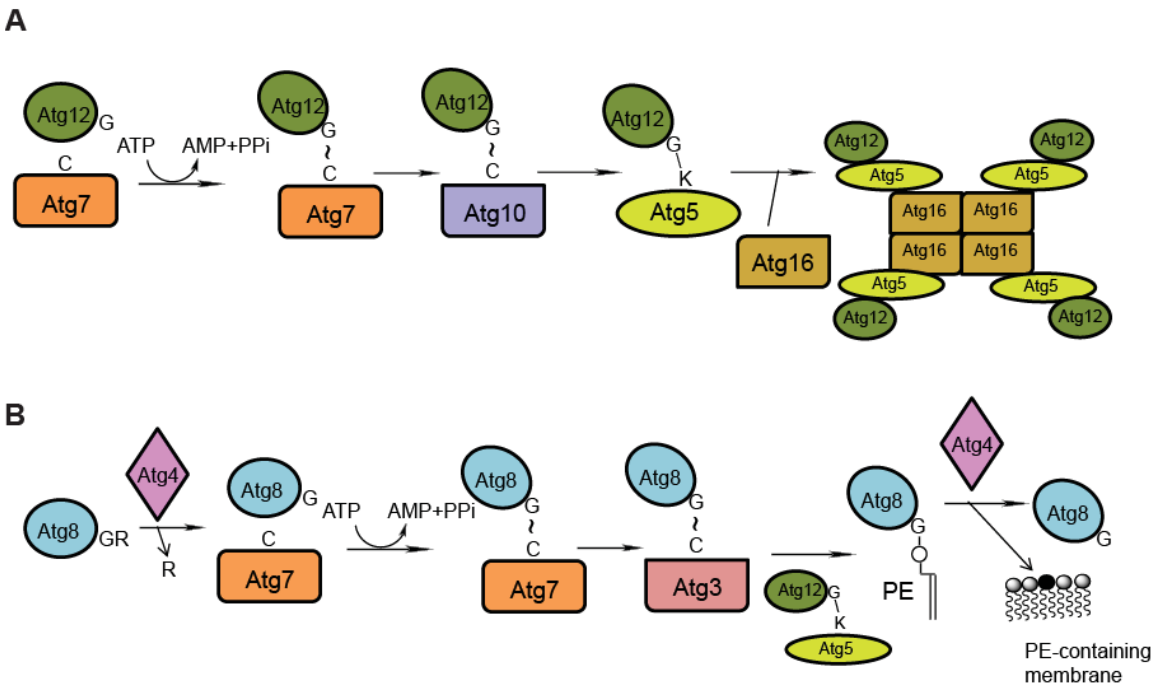


Figure 7. Schematic models of ubiquitin-like conjugation system for autophagosome formation in yeast.

(A) Atg-12-Atg5 conjugation system. (B) Atg8 conjugation system. PE, phosphatidylethanolamine.

5.1.4 Function

Studies have shown that macroautophagy is a tightly-regulated process that plays an important part in many physiological and pathological processes [174, 187, 188]. One role is to maintain cellular homeostasis, since autophagic degradation of cellular constituents efficiently recycles essential nutrients to sustain basic biological processes. Macroautophagy is required for normal development, in processes such as cellular differentiation, tissue remodeling and growth control. It is also used as a defense mechanism for cell adaptation to adverse environments, e.g. response to starvation, anti-aging mechanisms, and innate immunity. Deregulation of autophagy has been proposed to play a role in various diseases, including cancer, cardiomyopathy, muscular diseases, and neurodegenerative disorders.

Many questions about the process and the mechanism remain to be elucidated. There is mTOR dependent macroautophagy under nutrient deprivation conditions as well as mTOR independent basal autophagy for housekeeping or during myotube differentiation [189]. Bypass of the essential Atg5/Atg7 in autophagosome formation and autophagy-mediated protein degradation has also been reported in mouse embryonic fibroblast (MEF) cells and embryonic tissues [190]. The role of macroautophagy in disease is not well characterized either. It may help to halt the progression of some diseases and plays a protective role against intracellular pathogens [191], whereas in other situations, it may actually contribute to the development of a disease.

5.2 Microautophagy

In microautophagy, the cytoplasm is engulfed directly at the lysosomal surface, by invagination, protusion, and/or septation of the lysosomal limiting membrane. This process forms intralysosomal, single-membrane vesicles that are degraded by the lysosome [174].

5.3 Chaperone-mediated autophagy

In chaperone-mediated autophagy (CMA), the substrates to be degraded in lysosomes are soluble cytosolic proteins and usually possess a KFERQ motif (Lys-Phe-Glu-Arg-Gln). They are associated with a protein complex containing a cytosolic chaperone protein, the heat shock cognate protein of 70 KDa (hsc70) and a group of co-chaperones, which together help the protein substrates target to the receptor, lysosome-associated membrane protein type 2a (LAMP-2A). Multiple molecules of single-span transmembrane protein LAMP-2A together with lysosomal hsc70 are organized at the lysosome membrane to mediate translocation of protein substrates into the lysosomal lumen [192]. CMA has been considered to control normal cell functions and to supply energy in respond of nutritional stress, by clearing long-lived proteins. Dysfunction of CMA has been related to a variety of diseases, such as lysosomal storage diseases, nephropathies and neurodegenerative disorders [192]. The reduced occurrence and/or diminished activity of CMA with aging are thought as a major aggravating factor in some diseases, particularly neurodegenerative disorders. Study of the changes in CMA under pathological conditions could provide a potential therapeutic strategy [193].

5.4 Selective autophagy

In addition to CMA, growing evidence suggests that selective removal of protein aggregates, organelles, and microbes in cells is mediated by a macroautophagy process. A cargo receptor/adapter is required to connect the substrate to be degraded and the autophagasome proteins [194]. In yeast, selective cargo transportation to lysosome is demonstrated by the biosythetic Cvt pathway in which Atg19 plays the role of cargo receptor. Atg19 binds to two lysosomal enzyme precursors, the α-mannosidase (Ams1) and the dodecameric aggregate of aminopeptidase 1 (Ape1) as well as a critical autophagosome formation protein Atg8, forming Cvt complex. The complex is transported by Atg11 to PAS

where the double membrane is recruited and surrounds the complex [195]. Also, the selective disposal of peroxisomes requires autophagic receptor Atg30 in *Pichia pastoris*, and the removal of damaged mitochondria in *S. cerevisiae* needs Atg32 [196]. In mammalian cells, macroautophagy was thought to be a fairly unselective process. However, recent reports have identified several selective autophagy processes such as aggrephagy that removes protein inclusions called aggrosomes, pexophagy, mitophagy, reticulophagy and ribophagy dealing with different organelles, corresponding to peroxisomes, mitochondria, surplus ER and ribosomes respectively, lipophagy for oxidized lipid degradation, and xenophagy that eliminates intracellular pathogens like bacteria and viruses [194, 197].

Cargo receptors have been proposed to play a key role in each selective autophagy process. Sequestosome protein-1 (SQSTM1), also widely known as ubiquitin-binding protein p62, has been considered a cargo receptor for autophagic degradation of ubiquitinated protein substrates. SQSTM1/p62 is addionally a participant in Parkin-mediated mitophagy and in both vacuolar and cytosolic bacteria removal in which SQSTM1/p62 serves as a cooperator of calcium binding and coiled-coil domain 2 (CALCOCO2) also called antigen nuclear dot 52 kDa protein (NDP52) [194]. The much less studied next to BRCA1 gene 1 protein (NBR1) has been identified as a cargo receptor working as a SQSTM1/p62 partner in disposal of misfolded proteins [198]. The BCL2/adenovirus E1B 19 kDa protein-interacting protein 3-like (BNIP3L) which is also known as NIP3-like protein X (Nix) has been reported as a selective autophagy receptor for the clearance of mitochondria in a ubiquitin independent manner [199]. Expanding knowledge of selective autophagy would assist study and discovery of drug targets for various diseases like cancer, neurodegenerative diseases, diabetes and infections.

6. Atg8 family

The most conserved Atg protein, Atg8 has been suggested to have six mammalian orthologs. All Atg8 orthologs have the conserved glycine at their C-termini where PE modification, a key step in autophagosome formation, occurs. Based on homology of amino acid sequences, the Atg8 family has been divided into two subgroups [200]. Microtubule-associated protein 1A/1B-light chain 3 (MAP1-LC3) A, B and C have been grouped into the LC3 subfamily in which LC3A has two alternative splicing isoforms, a and b. γ-aminobutyric acid (GABA)$_A$ receptor-associated protein (GABARAP), GABARAP-like 1 (GABARAPL1), GABARAP-like 2 (GABARAPL2, also called golgi-associated ATPase enhancer of 16 kDa, GATE-16) have been placed in the GABARAP subfamily. A recent report proposed that the LC3 subfamily is more involved in isolation membrane elongation while the GABARAP subfamily plays a crucial part in autophagosome maturation [200]. The mouse Atg4b, a mammalian ortholog of yeast Atg4, has been found to act on the C termini of Atg8 orthologs using its active-site cysteine. Although the amino acid sequences of these Atg8 orthologs differ from one another by as much as 70%, their affinities for Atg4b were roughly comparable in competition experiments [201].

6.1 Microtubule-associated protein 1A/1B-light chain 3 (MAP1-LC3 or LC3)

LC3 was the first mammalian ortholog of Atg8 to be identified, in rat. There are four human LC3 proteins varying from 121 to 147 residues in length, which includes LC3A with two splicing isoforms, LC3B and LC3C. To date, only LC3B, the 125-residue protein has gained extensive attention and is commonly used to monitor autophagy since the cytosolic LC3 (LC3-I) can be lipidated during autophagy to LC3-PE (LC3-II) and incorporated into autophagosome membranes [202]. The two forms of LC3 are detectable by immunoblotting or immunofluorescence methods.

6.2 γ-Aminobutyric acid (GABA) $_A$ receptor-associated protein (GABARAP)

Type-A receptors for the neurotransmitter GABA (gamma-aminobutyric acid) are ligand-gated chloride channels that mediate inhibitory neurotransmission. Each subunit of the pentameric receptor protein has a large intracellular loop that is phosphorylated [203] and may be required for sub-cellular targeting and membrane clustering of the receptor, perhaps by anchoring the receptor to the cytoskeleton [204, 205].

Using the intracellular loop of the $GABA_A$ receptor γ-2S subunit (GABRG2) as bait, Wang *et al*. [206] identified a cDNA encoding GABA receptor assocated protein (GABARAP) by yeast two-hybrid screens of a fetal brain cDNA library and an adult brain cDNA library. Sequence analysis predicted that the 117-amino acid, 13.9-kDa GABARAP protein contains a basic N terminus and an acidic C terminus, with pI of 9.6. *In vitro* binding assays showed that the N-terminal 21 amino acids of GABARAP formed an α-helix that interacted with tubulin. Immunoprecipitation and immunohistochemical analysis in rat brain tissues demonstrated association and colocalization of GABARAP and $GABA_A$ receptors. Northern blot analysis detected a 0.9-kb GABARAP transcript in all tissues tested, namely heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas. Western blot analysis also showed a ubiquitous presence of GABARAP, suggesting that GABARAP is also involved in biologic events other than interaction with $GABA_A$ receptors.

As one of the mammalian orthologs of yeast Atg8, GABARAP is about 30% identical to LC3 [206]. Kebeya *et al*. [207] demonstrated that GABARAP also generates a form II state, possibly by PE conjugation similarly to what happens to LC3 during autophagy.

6.3 GABARAP-like 1 (GABARAPL1)

GABARAP-like 1 (GABARAPL1) was also named previously as glandular epithelial cell protein 1 (GEC1) or murine Atg8L/Apg8L.

Through a differential screen of estrogen-regulated sequences expressed in quiescent guinea pig endometrial glandular epithelial cells (GEC), Pellerin *et al*. [208] cloned guinea pig Gabarapl1, naming it, after the cell they used, Gec1. Using the guinea pig sequence as probe, Vernier-Magnin *et al*. [209] cloned human GABARAPL1 which they called GEC1, from a placenta cDNA library. GABARAPL1 is a 117-amino acid protein with the same size of GABARAP and 87% identity.

By Northern blot analysis of human tissues, a 2.3-kb GABARAPL1 transcript was detected in all tissues examined, with highest levels in brain, heart, peripheral blood leukocytes, liver, kidney, placenta and skeletal muscle. Moderate expression was found in pancreas, prostate, testis, ovary, lung, spleen, and colon, and low expression was found in thymus and small intestine. In addition, a 4.2-kb transcript was found in low abundance in ovary, peripheral blood leukocytes, and liver, as well as a 1.7-kb transcript in placenta at low levels [210]. Northern blot analysis of mouse tissues in the same report showed high levels of a 1.8-kb transcript in brain and kidney, and high levels of a 1.3-kb transcript in testis and heart [210]. A recent report has confirmed that GABARAPL1 associates with autophagic vesicles [211] as do other Atg8 family members. Moreover, GABARAPL1 has been shown in cultured cells, to cooperate with BNIP3L/Nix in selective autophagic disposal of damaged mitochondria [199].

6.4 GABARAP-like 2

GABARAP-like 2 (GABARAPL2) was first identified as a Golgi-associated ATPase enhancer of 16 kDa (GATE-16) interacting with N-ethylmaleimide

sensitive fusion protein (NSF) and Golgi v-snare [212]. Later, GABARAPL2 was shown to be present in autophagosomes induced by starvation [207].

6.5 Atg8 family interacting motif and the docking site

Several analyses have suggested that the interaction between cargo receptors and Atg8 family proteins is mediated by specific and conserved sequences on the cargo receptors, a so-called Atg8 family interacting motif (AIM) [213] or LC3 interacting region (LIR) [167]. The core sequence of an typical AIM is WxxL-like sequence [213] (Figure 8). Based on current knowledge of those reported Atg8 family interacting proteins, the consensus sequence of AIM could be refined as X_{-3}-X_{-2}-X_{-1}-W/F/Y-X_1-X_2-L/I/V, where acidic residues usually occupy X_1 and/or X_2 and at least one of X_{-3}-X_{-2}-X_{-1} positions [194].

	AIM sequence	Binding partner
Atg19	ALTWEEL	Atg8
Atg32	SGSWQAIQP	Atg8
Atg3	VGDWEDLQD	Atg8
Atg4B	TLTYDTLRF	LC3
p62	DDDWTHLSS	LC3
NBR1	SEDYIIILP	LC3
CRT	EDDWDFLPP	GABARAP
CHC	TPDWIFLLR	GABARAP
Nix	NSSWVELPM	GABARAP
(PSSM)	DDDWVFVPM	GABARAP

Adapted from Noda NN, Ohsumi Y and Inagaki F. FEBS Lett. (2010)

Figure 8. Sequence alignment of reported AIMs and their interacting Atg8 family proteins [213].

p62, Sequestosome protein-1 (SQSTM1); NBR1, Next to BRCA1 gene 1 protein; CRT, calreticulin; CHC, clathrin heavy chain; PSSM, sequence position-specific scoring matrix determined by phage display screening.

A conserved surface, LIR docking site (LDS) on Atg8-family proteins is used for interaction with the hydrophobic AIM. The LDS on Atg8-family proteins consists of two hydrophobic pockets, W-site and L-site, formed by the side-chains of two groups of hydrophobic residues [214]. The aromatic residue is required for the W-site of LDS and the large hydrophobic residue is required for the L-site [194]. Mutations in LDS or AIM impair the binding of Atg8-family proteins to AIM-containing cargo receptors and hence disrupt the cargo transport [167, 198, 199, 214].

RESEARCH OBJECTIVE

Since the first report on the cloning of human Stbd1 in 1998, few studies have addressed this novel protein except for a couple of bioinformatic analyses listing it as one of the CBM20-containing proteins. Through different yeast two-hybrid screens carried out previously in our laboratory (Skurat, A.V. unpublished data), Stbd1 showed up as potentially interacting with several proteins involved in glycogen metabolism, e.g. GS, GBE and DBE/AGL. When a truncated form of human Stbd1 was used as bait for a yeast two-hybrid screen, the identified targets included two autophagy related proteins. From this background, we sought to understand the role of Stbd1 in the context of glycogen metabolism and whether it had any relevance to an autophagy-like pathway for lysosomal degradation of glycogen.

EXPERIMENTAL PROCEDURES

1. Yeast two hybrid screen

The yeast *Saccharomyces cerevisiae* strain PJ69-4A transformed with the previously constructed bait plasmid pGBDU-GenethC containing the cDNA that encodes 179-358 of human Stbd1 was subsequently transformed with a human skeletal muscle cDNA library HL4010 (Clontech) in pGAD vector. Positive colonies were selected from synthetic medium lacking uracil, leucine, and adenine (Sc-Ura-Leu-Ade) and re-plated on synthetic medium without uracil, leucine, and histidine (Sc-Ura-Leu-His) but with 5 mM 3-amino-1,2,4-triazole (3-AT). The library plasmids selected from the second screen were rescued in *E. coli* RRI cells via electroporation and screened by M9 medium deficient in leucine. The rescued plasmids were analyzed by yeast two-hybrid tests and DNA sequencing. The purified pGAD plasmids were transformed back into yeast with pGBD plasmids and quantitative assay of β-galactosidase activity was carried out to evaluate interactions [215]. (Figure 9)

1.1 Yeast transformation (lithium-acetate method)

Five ml of overnight culture of yeast was aliquotted and pelleted down in 1.5 ml microcentrifugation tubes at 18,000 g (Beckman counter microfuge 22R centriguge) for 5s. After removal of supernatant, 240 µl of 50% (w/v) PEG3350, 36 µl of 1 M LiAc, pH7.5, 5 µl of 10 mg/ml single strand carrier DNA (ssDNA), 0.1-5 µg of plasmid DNA and 65 µl of sterile H_2O were added to the pellet and vortexed for 1min to mix well. The mixture was then incubated at 30°C for 5min followed by 20min incubation at 42°C. After 10s centrifugation at 18,000 g, the pellet was resuspended in 0.2 ml of sterile H_2O and plated on selective plates. The plates were incubated at 30°C. For cDNA library transformation, yeast cells were harvested by centrifugation at 18,000 g for 5min when cell titers reached

$2x10^7$cells/ml and washed with sterile H_2O. To the pellet, 10 volumes of the above reagents and DNAs were added.

1.2 Isolation of yeast DNA

Ten ml of overnight yeast culture was centrifuged and the pellet was resuspended in 0.2 ml of yeast lysis solution containing 2% Triton X-100, 1% SDS, 100 mM NaCl, 100 mM Tris-HCl, pH8.0, 1 mM EDTA, 0.2 ml of phenol: chloroform ($CHCl_3$): Isoamyl alcohol (IAA) (25:24:1) and 0.3g of acid-washed glass beads (425-600 µm). The mixture was vortexed for 2 min followed by centrifugation at 18,000 g for 5 min. Supernatant, 180 µl, was transferred to a clean microcentrifugation tube, and 20 µl of 3M NaAc and 0.5 ml of cold ethanol were added. After 10 min centrifugation at 18,000 g, the pellet was washed with 0.2 ml of 70% ethanol, dried by vacuum for 5 min and resuspended in about 50 µl of TE buffer containing 10 mM Tris and 1 mM EDTA.

1.3 Rescue of plasmid in *E. coli* RRI cells by electroporation

Thawed on ice, 40 µl of *E. coli* RRI competent cells were added 2 µl of DNA isolated from yeast. The mixture was transferred to an electroporation cuvette with 0.1 cm electrode gap (BioRad) and the voltage was set to 2.5 kV. After exposure to high voltage, 1 ml of LB was added quickly. The mixture was transferred to a clean microcentrifugation tube and incubated in shaker at 37°C, 200 rpm for 1h. The pellet was resuspended in 0.8-1 ml of 10% glycerol and 100-150 µl of the resuspended cells were plated on M9 medium lacking leucine (M9-Sc-Leu) with 0.1 mg/ml ampicillin, 1 mM thiamine-HCl and 0.01 mg/ml proline. The plates were incubated at 37°C for at least 16h.

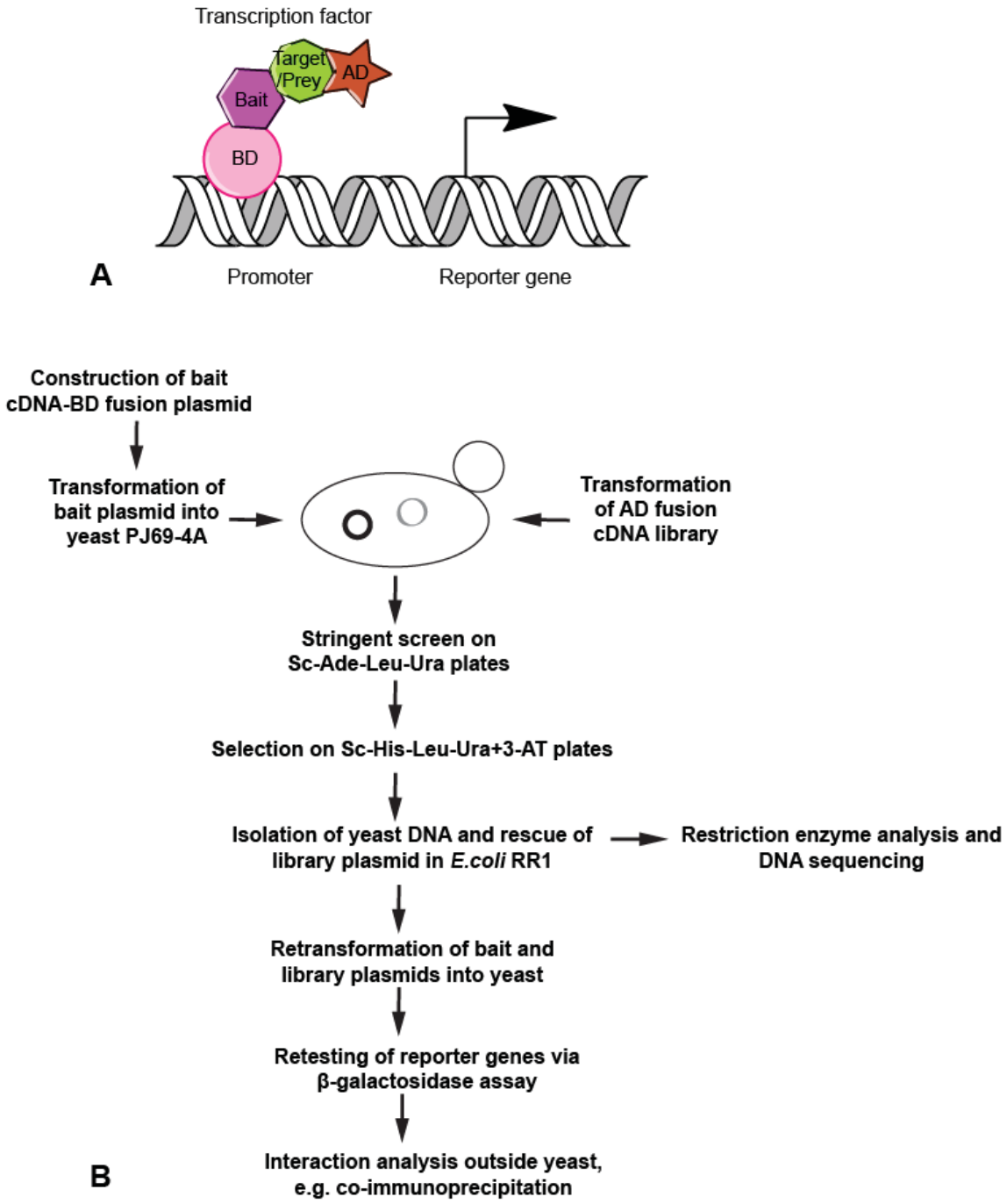


Figure 9. Yeast two hybrid screen.

(A) Principle of the system. BD: DNA-binding domain; AD: transcriptional activation domain. (B) Flow chart of the system.

1.4 β-galactosidase activity assay

Individual yeast colonies were inoculated into appropriated synthetic medium containing glucose and incubated in a shaker at 200rpm at 30°C till OD_{600} =~1. On ice, 0.1 ml of overnight culture or H_2O as blank was mixed in 1.5 ml microcentrifugation tube with 0.7 ml of Z-buffer (60 mM $Na_2HPO_4.7H_2O$, 40 mM $NaH_2PO_4.H_2O$, 10 mM KCl, 1 mM $MgSO_4.7H2O$, pH7.0) with 50 mM of 2-mercaptoethanol, 50 µl of chloroform ($CHCl_3$) and 50 µl of 0.1% SDS. The mixture was vortexed for 2min and 0.16 ml of freshly made O-nitrophenyl-β-D-galactopyranoside (ONPG) solution containing 4mg/ml ONPG, 60 mM $Na_2HPO_4.7H_2O$, 40 mM $NaH_2PO_4.H_2O$, pH7.0 was added. After 60min incubation at 30°C, the reaction was quenched by adding 0.4 ml of 1M Na_2CO_3. Cell debris were pellet down after 10min of centrifugation at the 18,000 g and 1 ml of supernatant was transferred into clean cuvette to read OD_{420}. The linear range is around 0.2-0.8. Meanwhile, 0.1 ml of overnight yeast culture kept on ice was mixed with 0.9 ml of H_2O to read OD_{600}, using H_2O as blank. The β-galactosidase activity was calculated as following: β-galactosidase activity units=1000 x OD_{420} / (t x V x OD_{600}) in which t= incubation time (min), V= volume (ml) of culture added to Z-buffer.

2. Plasmid construction

The cDNA corresponding to the coding region of human Stbd1 (1076bp) was amplified by polymerase chain reaction (PCR) (Initial denaturation at 95°C for 5 min; 30 cycles of denaturation at 94°C for 1 min, annealing at 65°C for 1 min, extension at 72°C for 2 min; and the last extension at 72°C for 10 min). The primers (forward: 5'-G↓GA TCC ATG GGC GCC GTC TGG TCC GCC C-3' and reverse: 5'-G↓AA TTC TCA AGC GTA ATC TGG AAC ATC GTA TGG GTA GTG AAT CCC CCA CCA TGC GTG A-3') were designed to introduce restriction sites *Bam*HI and *Eco*RI as underlined, and the sequence of the HA tag (YPYDVPDYA) expressed at the C-terminus of the protein as curved underlined. The TA cloning

kit was used to insert the amplified DNA fragment into pCR2.1 vector. When cDNA sequences of human Stbd1 coding region were verified by the DNA sequencing Core Facility at the Indiana University School of Medicine, the cDNA was subcloned to pcDNA3 vector for mammalian expression (Figure 10), and pET28a/pET32a and pGEX vectors for *E. coli* expression, via restriction digestion at 37°C and ligation at 16°C. The cDNA corresponding to the coding region of mouse Stbd1 (1016bp) was also amplified and cloned with the same strategy (Figure 11). The primers are forward: 5'-G↓GA TCC ATG GGC GCC GTC TGG TCA GCC C-3' and reverse: 5'- G↓AA TTC TCA AGC GTA ATC TGG AAC ATC GTA TGG GTA GTG AAT CCC CCA CCA CCC ATG A-3'.

The cloning of GABARAP and GABARAPL1 followed the similar procedure (Figure 12). The primers for GABARAP (forward: GC↓G GCC GCT ATG AAG TTC GTG TAC AAA and reverse: T↓CT AGA TCA CAG ACC GTA GAC ACT TTC GTC AC) were designed to introduce restriction sites NotI and XbaI as underlined. The primers for GABARAPL1 (forward: 5'-G↓ AAT TCA ATG AAG TTC CAG TAC AAG GAG GAC CAT-3' and reverse: 5'-T↓CT AGA TCA TTT CCC ATA GAC ACT CTC ATC ACT G-3') were designed to introduce restriction sites *EcoR*I and *Xba*I as underlined. The cDNA of GABARAP and GABARAPL1 coding region were subcloned by restriction digestion and ligation with the pFlagCMV vector for mammalian expression, and pET32a and pGEX vectors for *E. coli* expression.

3. Ligation, transformation and plasmid preparation

PCR product and pCR2.1 vector (Invitrogen), or DNA fragment and vector plasmid yielded by restriction digestion were mixed with deionized sterile H_2O, ligation buffer and T4 ligase (New England Biolabs) and incubated at 16°C for 16-20h.

Ligation product, 5 µl or 50ng of plasmid was added to 100 µl competent cell of *E. coli* DH5α thawed on ice. The mixture was set on ice for 15min followed by heat shock at 42°C for 45s. LB, 500 µl was added and the mixture was incubated at 37°C, shaking at 200 rpm for 1h. The culture was pelleted down by quick centrifugation and 300 µl of the supernatant was removed. The pellet was resuspended in the rest supernatant and 50 µl or 100 µl of the suspension was plated on LB plates with proper antibiotics. The plates were incubated at 37°C for at least 16h.

Plasmids were prepared by BioRad miniprep kit following factory instruction.

4. Mutagenesis

Mammalian expression vectors containing HA tagged hStbd1 (hStbd1-HA) and different truncation mutants thereof (ΔN24-HA, ΔN90-HA, ΔC96-HA) for mammalian expression were made by PCR amplification of a human cDNA with addition of an HA tag at the C-terminus. The products were subcloned into *Bam*HI/*Eco*RI sites of the pCDNA3 vector (Figure 10 and 14). Plasmids with point mutations or short deletions in the carbohydrate binding domain (W293G-HA and W293L-HA) and the Atg8-family interacting motif (Δ198-222-HA, W203A-HA, V206A-HA, W212A-HA, V215A, W203A+V206A-HA, W212A+V215A-HA) were constructed by site-directed mutation using pCDNA3-hStbd1-HA as template (Figure 13 and 14). The primers containing the mutated sequences were designed using the online service of PrimerX (http://www.bioinformatics.org/primerx/) and were synthesized by Invitrogen. The *Pfu* Turbo DNA polymerase (Stratagene) was used for the PCR. PCR cycling condition: initial denauration was at 95°C for 5 min, denaruation was at 95°C for 1 min, 20 cycles including denaturation at 95°C for 1 min, annealing at 60°C for 1 min and extension at 68°C for 12 min.
The PCR product was digested with *Dpn* I (New England Biolabs) at 37°C for 2-3h to remove the parental DNA. The PCR product was transfected into

competent cells to generate pcDNA3-W293G-HA and pcDNA3-W293L-HA,. His-tagged or GST-fusion hStbd1 (human Stbd1) and mStbd1 (mouse Stbd1) vectors for bacterial expression were constructed by subcloning into pET28a or pGEX vectors.

Sequences of all constructs were verified by the DNA sequencing Core Facility at the Indiana University School of Medicine.

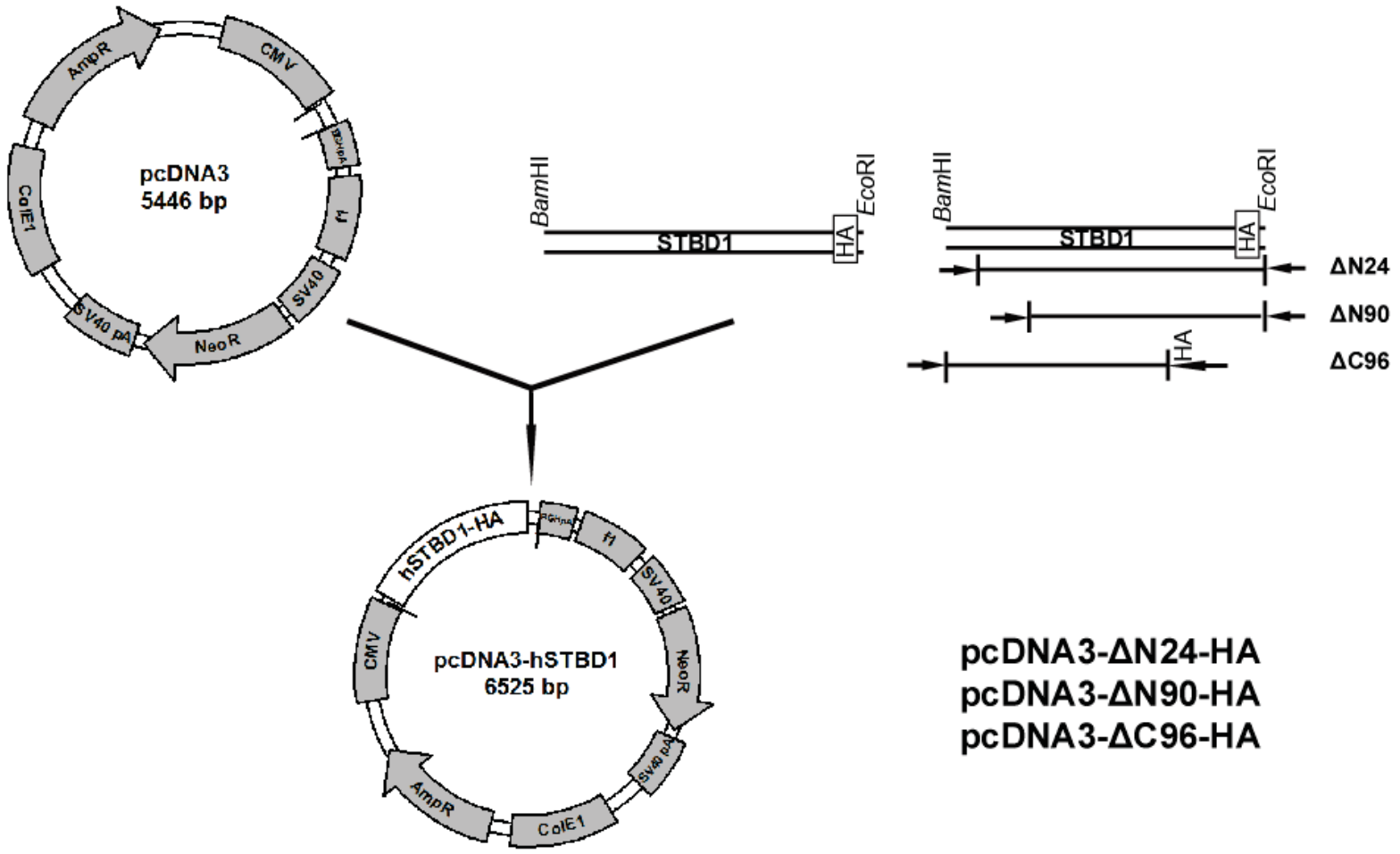


Figure 10. Plasmid construction map of mammalian expression plasmids containing human STBD1 with C-terminal HA tag in different lengths.

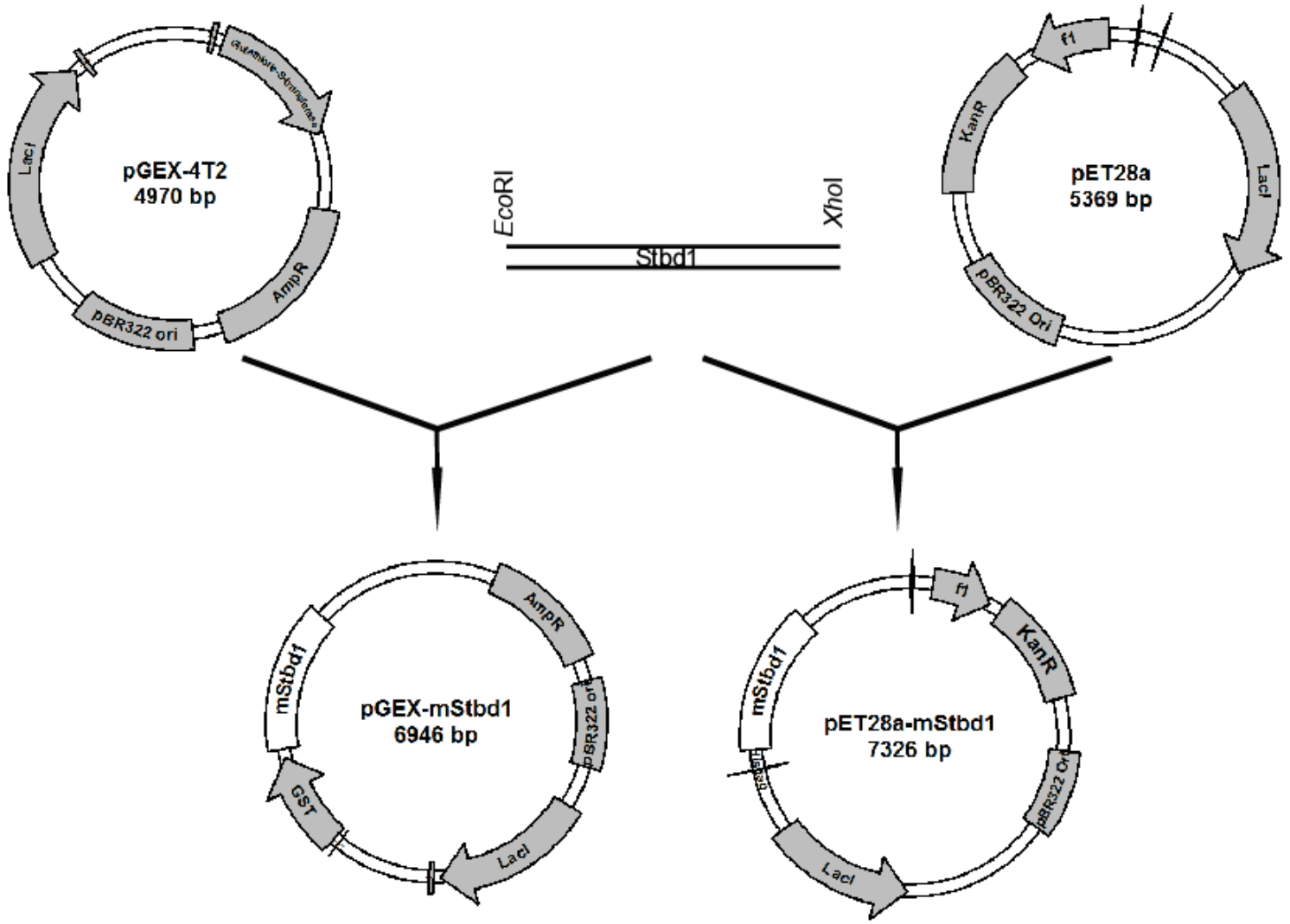


Figure 11. Plasmid construction map of *E. coli* expression plasmids containing mouse Stbd1.

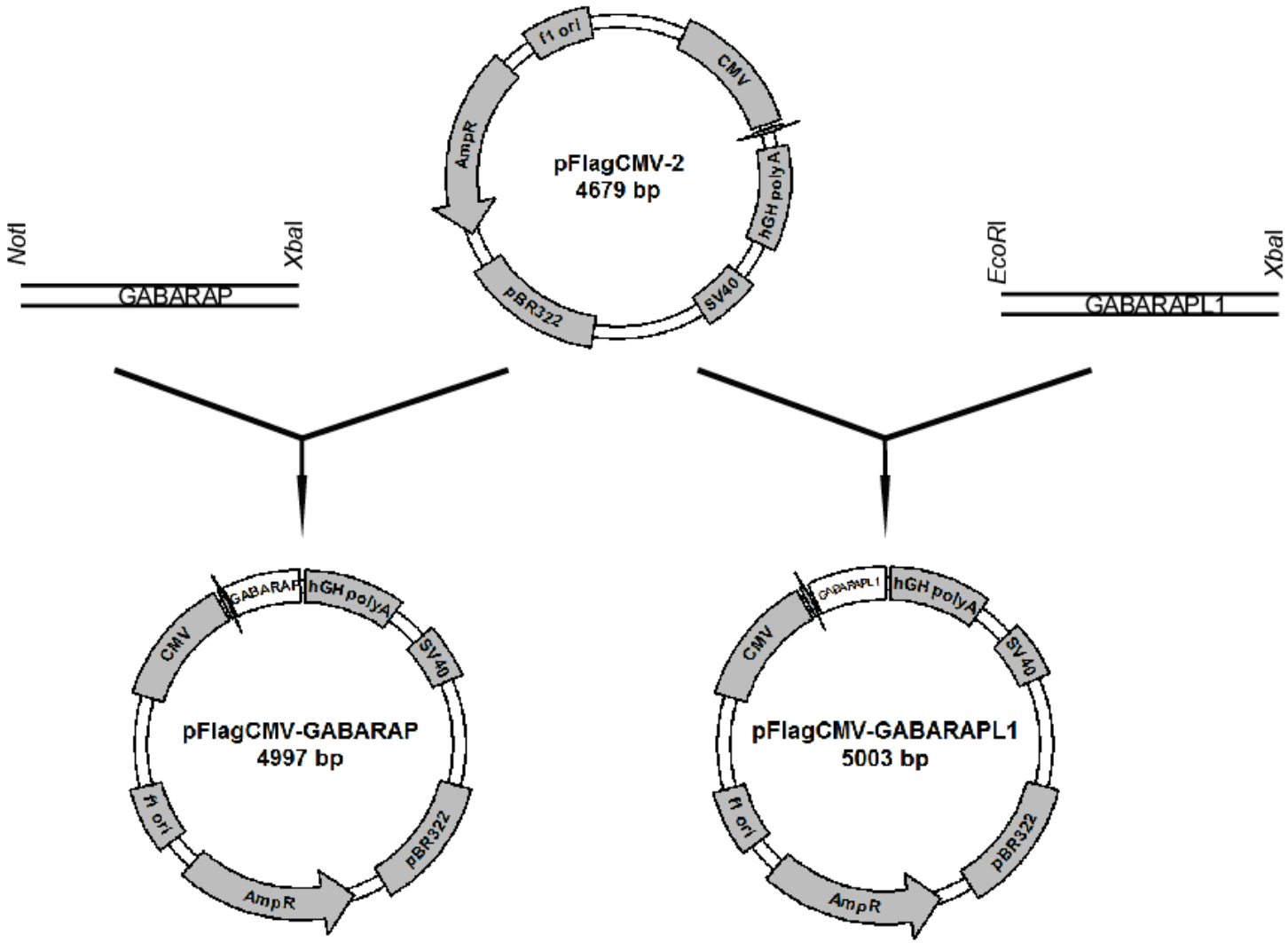
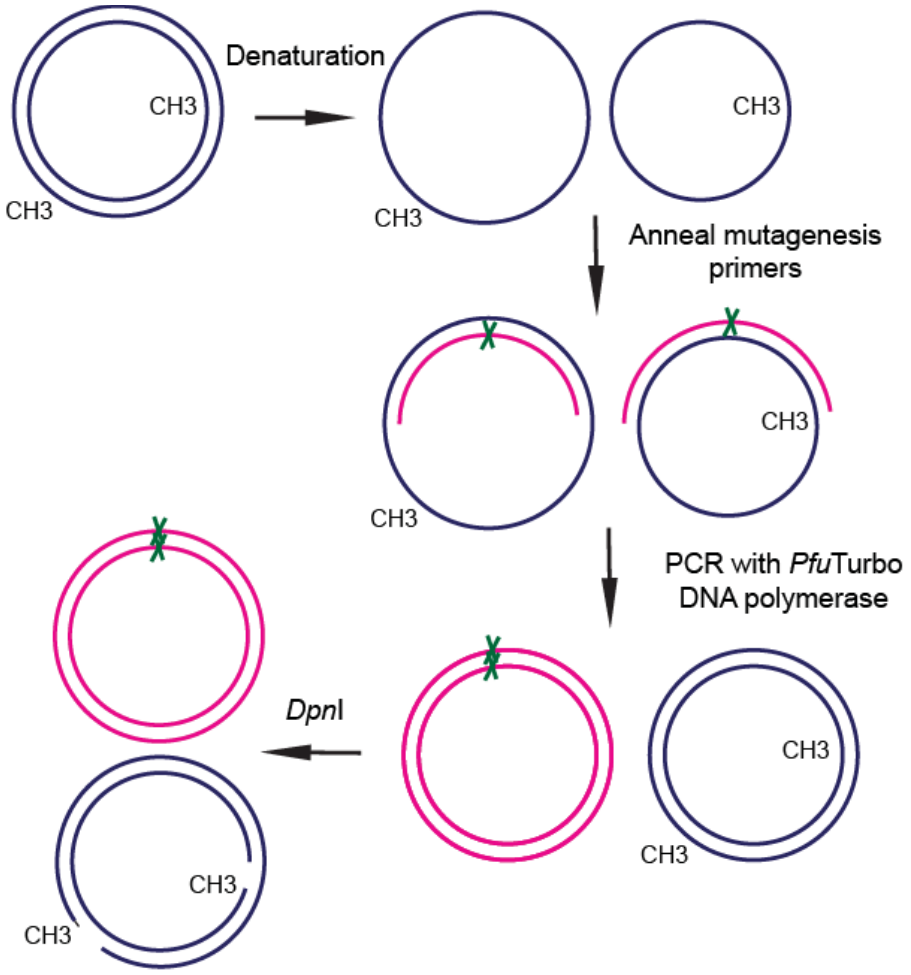


Figure 12. Plasmid construction map of mammalian expression plasmids containing human GABARAP or GABARAPL1.



W293G	**TGG-->GGC**
W293L	**TGG-->CTG**
Δ198-222	**GACCGG GCTCCA**
W203A	**TGG-->GCC**
V206A	**GTG-->GCC**
W212A	**TGG-->GCC**
V215A	**GGG-->GCC**
W203A+V206A	**TGG-->GCC, GTG-->GCC**
W212A+V215A	**GGG-->GCC, GGG-->GCC**

Figure 13. Site-directed mutegenesis.

Point mutations of conserved trypsine residue within CBM20 into glycine or leucine, W293G and W293L). Deletion and single or double point mutations of potential ATG8 family interacting motif (AIM), Δ198-222, W203A, V206A, W212A, V215A, W203A+V206A and W212A+V215A).

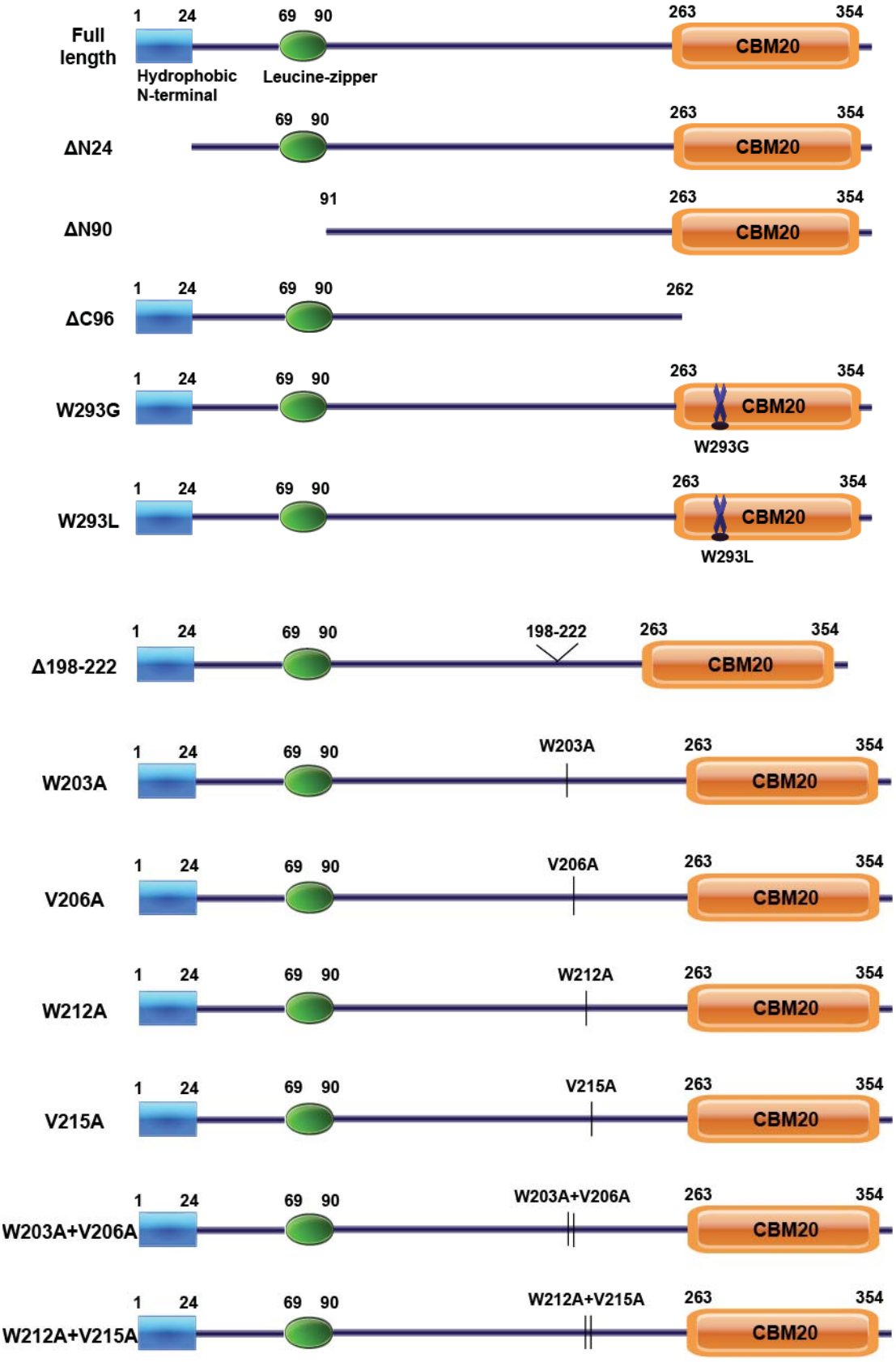

Full length
1
24
69
90
263
354
CBM20
Hydrophobic N-terminal
Leucine-zipper
ΔN24
ΔN90
91
ΔC96
262
W293G
W293L
Δ198-222
198-222
W203A
V206A
W212A
V215A
W203A+V206A
W212A+V215A

Figure 14. Schematic architecture of human Stbd1 and mutants.
(A) Full length human Stbd1; Deletion of the N-terminal hydrophobic segment (ΔN24-HA); Deletion of the N-terminal hydrophobic segment and the putative leucine zipper (ΔN90-HA); Deletion of the C-terminal carbohydrate binding domain (ΔC96-HA); Mutation of conserved Trp of CBM20 domain to Gly (W293G); Mutation of conserved Trp of CBM20 domain to Leu (W293L); Deletion of region including putative Atg8 family interacting motif (AIM) /LC3 interacting region (LIR); Mutation of conserved Trp or Val into Ala (W203A, V206A) within the first potential AIM; Mutation of conserved Trp or Val into Ala (W212A, V215A) within the second potential AIM; Double mutation of both conserved Trp and Val into Ala (W203A+V206A) within the first potential AIM; Double mutation of both conserved Trp and Val into Ala (W212A+V215A) within the second potential AIM.

5. Expression and purification of recombinant proteins

E. coli BL21 (DE3) competent cells transformed with pET28a-hStbd1, pET28a-mStbd1 and pGEX-mStbd1 vectors were grown at 37°C till the OD_{600} reached 0.4. Protein expression was induced by 0.4 mM isopropyl-β-d-thiogalactoside (IPTG) at 18°C overnight. The His-tagged hStbd1 and mStbd1 were purified by Ni^{2+}–NTA agarose chromatography (Qiagen). The proteins were eluted stepwise with 40 mM to 200 mM imidazole. hStbd1 and mStbd1 were eluted primarily in fractions with 100 and 200 mM imidazole. The eluted fractions were dialyzed against buffer containing 50 mM Tris HCl, pH 7.5, 150 mM NaCl and 5% 2-mercaptoethanol and stored at −80°C with 15% glycerol. The His-tagged mStbd1 was dialyzed against PBS, pH 7.4 and to Affi-Gel 15 for antibody affinity purification. The GST-fusion mStbd1 for antibody generation was purified using glutathione Agarose (Sigma). The recombinant protein was eluted with 10 mM glutathione and dialyzed against phosphate-buffered saline (PBS).

6. Glycogen purification and polysaccharide binding assay

Glycogen was purified from mouse muscle as previously described [216]. Briefly, tissues were boiled in 30% KOH and lipids were removed by the methanol/chloroform extraction. Glycogen precipitated in cold ethanol and redissoved in H_2O was treated with 10% trichloroacetic acid (TCA). After centrifugation, glycogen recovered from the supernatant by ethanol precipitation was subjected to extensive dialysis.

Recombinant hStbd1 was diluted in 50 mM Tris HCl pH7.5, 150 mM NaCl, 0.1% 2-mercaptoethanol, 0.1% TritonX-100 and subjected to 10,000 g centrifugation at 4°C for 20 min. The supernatant was mixed with the supernatant of glycogen/amylopectin centrifuged under same conditions. The final concentration of protein was 2.5 μg/ml and glycogen/amylopectin was 0.25 mg/ml. Stbd1 without polysaccharide was used as negative control. The

samples were mixed well, incubated at 4°C for 1h, and then subjected to ultracentrifugation at 100,000 g at 4°C for 90 min. The pellets were resuspended by sonication. The supernatant and pellets were subjected to western blotting for hStbd1.

7. Antibodies

Rabbit polycolonal antibodies were generated against GST-mStbd1 by Cocalico Biologicals. The antibody was Affinity purified by Affi-Gel 15 (Bio-Rad) coupled with His-tagged mStbd. Rabbit polycolonal anti-HA epitope tag antibodies were from Rockland (Gilbertsville). Mouse monoclonal anti-HA epitope tag antibodies were from Covance (Emeryville). Mouse monoclonal anti-Flag M2 antibodies were from Sigma. Rabbit monoclonal Syntaxin 6 and glycogen synthase antibodies were from Cell signaling. Mouse monoclonal glycogen synthase antibodies were from Invitrogen. Rabbit monoclonal GS antibodies were from Cell signaling. Rabbit polyclonal anti-hStbd1 antibodies (anti-GENX-3414) were from Protein Tech Group. Rabbit polycolonal GABARAPL1antibodies were from Protein Tech Group. Mouse monoclonal LC3 antibodies were from NanoTools.

8. Cell culture and transfections

COS M9 cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Mediatech or Sigma) with 25 mM glucose and 10% fetal bovine serum (FBS)(Atalanta Biologicals, Lawrenceville). Transfections were performed with Fugene 6 (Roche) following the manufacturer's instructions. Six-well plate or 35 mm plates were used for immunoblotting or immunoprecipitation, and for immunofluorescence microscopy, 12-well plates were used with cells growing on coverslids at the bottom of each well. Glucose starved COS M9 cells were incubated in DMEM without glucose and supplemented with sodium pyruvate (Invitrogen) and 10% FBS. The normal "fed" condition had 25 mM D-glucose (Sigma).

Rat1Neo5 and Rat1WT4 cells [217] were grown in DMEM with 25 mM glucose and supplemented with 10% fetal bovine serum and 0.25 mg/ml geneticin. Rat1WT4 is a stably transfected cell line over-expressing wild type rabbit muscle glycogen synthase. Rat1Neo5 cells are a control line transfected with vector. FL83B cells (ATCC) were grown in F-12K medium (ATCC) and supplemented with 10% FBS. C2C12, Hepa1c1c, HepG2, HII4EC3 mouse embryonic fibroblasts (MEF) were grown in DMEM with 25 mM glucose with 10% FBS. C2C12 cells were differentiated in DMEM with 25 mM glucose and supplemented with 1% bovine calf serum. All cells were incubated at 37°C with 5% $CO_{2.}$

9. Preparation of tissue and cell extracts and immunoblotting

MGSKO [95] and LGSKO mice [93] have been described previously. They carry disruptions of the glycogen synthase genes *Gys1* or *Gys2* respectively. The exercise protocol followed a previous study [218]. Mice were sacrificed by cervical dislocation. Tissues were rapidly frozen in liquid nitrogen and stored at −80°C. Frozen tissues were powdered and homogenized in buffer containing 50 mM Tris HCl (pH 7.8), 10 mM EDTA, 2 mM EGTA, 0.1 mM *N-p*-tosyl-l-lysine chloromethyl ketone (TLCK), 2 mM benzamidine, 0.5 mM phenylmethylsulfonyl fluoride (PMSF), 50 mM 2-mercaptoethanol and 10 µg/ml of leupeptin, with or without 0.5% Triton-X. Tissue homogenates were used for Western blotting analysis. For analysis of mStbd1 and glycogen fractionation, tissue homogenates were first subjected to low speed centrifugation 10,000 g at 4°C for 10 min to generate a low speed pellet (LSP) and the supernatant (LSS) was subjected to further centrifugation. Centrifugation at 100,000 g at 4°C for 90 min, generated a high speed supernatant (HSS) and glycogen was collected in the high speed pellet fraction (HSP). The HSP was resuspended in the starting volume of homogenization buffer. All fractions were analyzed by Western blotting with anti-mStbd1 antibodies.

Cultured cells were lysed in buffer containing 50 mM Tris HCl, 100 mM NaCl, pH 7.5, 0.5% TritonX-100, 1 mM PMSF, 0.1 mM TLCK, 1 mM benzamidine, 1 μg/ml of aprotinin, pepstatin and leupetin. The cell lysates were centrifuged at 10,000 g for 15 min at 4°C to pellet insoluble materials.

Protein concentration was determined by the Bradford method using BSA as standard [219]. Samples were subjected to 10% SDS-PAGE. Proteins were transferred to nitrocellulose membranes, blocked in Tris buffered saline (TBS) with 5% milk and incubated with antibodies diluted in TBS with 2% milk, followed by incubation with horseradish peroxidase (HRP)-conjugated secondary antibodies diluted in TBS with 2% milk and development by ECL western blotting substrate (Thermo Scientific).

10. Co-immunoprecipitation

COSM9 cells co-expressing C-terminal HA-tagged hStbd1and N-terminal Flag-tagged proteins like GABARAP, GABARAPL1 or laforin were lysed and centrifuged at 8,000 g for 15 min at 4°C. EZview red anti-Flag gel or EZview red anti-HA affinity gel (Sigma) was equilibrated with lysis buffer before use. The supernatant from cell lysates was mixed with affinity gel and incubated at 4°C overnight. The samples were centrifuged at 8,000 g for 30s to pellet the agarose. The agarose was washed by lysis buffer three times and loading buffer was added for SDS-PAGE and immunoblotting.

11. Immunofluorescence staining and microscopy

Cells were grown on glass coverslips 24 h or 48 h before fixation. Cells were fixed in PBS with 4% paraformaldehyde (PFA) for 3 x10min with 2min PBS wash in between, and then quenched and permeabilized in PBS with 100 mM glycine and 0.2% TritonX-100. Nonspecific binding sites were blocked with 5% bovine serum albumin (BSA) (Sigma) in PBS. The cells were incubated with primary

antibodies in PBS with 2% BSA at room temperature for 1-2h. Antibody dilutions were as follows: anti-mStbd1, 1:200; anti-HA, 1:1000; anti-FLAG, 1:1000; anti-GABARAPL1, 1:100; anti-LC3, 1:100; and anti-glycogen synthase, 1:200. Cells were then washed in PBS with 2% BSA and developed with secondary antibodies (1:400, Invitrogen), conjugated with either Texas Red or Alexa488 fluorophores. Nuclei were visualized by staining with 1 μg/ml Hoechst (Invitrogen). The specimens were washed in PBS and then milliQ H_2O before mounting. The slides were viewed on a Zeiss Axio Observer Z1 microscope with a Plan Apochromat ×63 oil immersion objective (Zeiss). Images were processed with Zeiss Axiovision 4.7.

12. Immunohistochemistry

Fixed mouse liver or muscle sections were deparaffinized by soaking in the order into xylene for 3x3min, 100% ethanol for 2x3min, 95% ethanol for 3min and 70% ethanol for 3min. The antigen was retrieved by soaking the specimens into 10 mM citrate buffer pH6.0 with 0.05% Tween 20 that heated to ~90°C for 20min till the temperature dropped to ~50°C and then cooled down to room temperature. The specimens were blocked in TBS with 1% BSA and 10% goat normal serum and incubated with anti-mStbd1 diluted in TBS with 1% BSA. Washed by TBS with 0.025% TritonX-100, the endogenous peroxidase of specimens was inactivated by TBS with 0.3% H_2O_2. HRP-conjugated secondary antibody was diluted in TBS with 1% BSA. The staining was developed with 3,3'-diaminobenzidine (DAB) substrate reagent (BD Biosciences) in the presence of HRP. Harris' hematoxylin was used for counterstaining to demonstrate nuclei. When immunohistochemical staining was followed by PAS staining, the counterstaining step was omitted. The samples were dehydrated before mounting by soaking in the order into 70% ethanol for 3min, 95% ethanol for 3min, 100% ethanol for 2x3min, xylene:ethanol (1:1) for 3min and xylene for 3x3min. The stained samples were mounted by xylene substitute mountant (Thermo Scientific) and viewed on a Leica DM3000 microscope (Leica).

13. Periodic acid-Schiff reagent (PAS) staining and microscopy

Fixed cells were rinsed in ddH_2O and oxidized with 0.5% periodic acid for 5 min, stained with Schiff reagent for 15 minutes and then counterstained in Harris' hematoxylin for 1 min. Negative controls without periodic acid treatment were processed at the same time to ensure staining specificity. When immunofluorescence staining was applied following by PAS staining, the counterstaining step was omitted. The method for the combined PAS and immunofluorescence staining was adapted from reference [220]. The PAS stained samples were viewed on a Leica DM3000 microscope (Leica). Images were processed with Leica Application Suite version 3.5.0. The samples with both PAS and immunofluorescence staining were viewed on a Zeiss Axio Observer Z1 microscope and images were processed with Zeiss Axiovision 4.7.

14. Glycogen assay

Sample and standard glycogen was dissolved in 0.3 mg/ml of amyloglucosidase in 200 mM NAOAc and digested for 3-4h (or overnight). Samples were diluted in H_2O after digestion. Glycogen was measured in triethanolamine-G6PDH-NADP buffer which included 300 mM triethanolamine, 4 mM $MgCl_2$, 2 µg/ml G6PDH, 0.9 mM β-NADP and 2 mM ATP. In 300 µl of triethanolamine-G6PDH-NADP buffer, 10 µl of proper diluted sample or standard glycogen was added and 100 µl of mixture was read at OD_{340} in 50 µl quartz cuvette. The triethanolamine-G6PDH-NADP buffer without adding glycogen was used as blank. Hexokinase was diluted in 3.2 M $(NH_4)_2SO_4$ to 1/2 and 1 µl was added to each 200 µl of sample or standard mixture. After 30 min incubation, samples were read at OD_{340} again and glycogen content (µg glucose/mg protein) can be calculated by standard curve or by the following formula (µmol glucose/g protein):
$(\Delta\ OD_{340}/6.22) \times 0.31\ ml \times (0.2\ ml/0.01\ ml) \times$ dilution factor/protein (g)

15. RNA isolation and quantitative PCR

RNA was isolated as previous described [93] using TRIzol reagent (Invitrogen) with manufactory instruction. Primers and probes were designed through Universal probe library tool (Roche Applied Science). Superscript III-RT kit (Invitrogen) was used for reverse transcription of isolated RNA. LightCycler 480 probes Master reagent and real time PCR system (Roche Applied Science) were used for the real time PCRs. The mRNA content was calculated as previously described [221] and the fold change was normalized by reference gene 18S

16. Statistical Analysis

The data are presented as mean ± standard error. Statistical significance was determined by unpaired Student's t test and significance was assigned at $p < 0.05$.

RESULTS

1. Association of Stbd1 with polysaccharides

Because of the presence of the CBM20 domain on Stbd1, our initial goal was to analyze the polysaccharides-binding ability of Stbd1 *in vitro* and *in vivo*.

1.1 Stbd1 binds glycogen and amylopectin *in vitro*

For the *in vitro* binding assay, recombinant human Stbd1 was produced in *E. coli* and purified. Polysaccharide binding was judged by co-sedimentation of Stbd1 with either glycogen or amylopectin during ultracentrifugation (100,000 g) (Figure 15A). The presence of Stbd1 was visualized by immunoblotting. Based on this assay, Stbd1 is capable of binding to both glycogen and amylopectin, but bound more effectively to amylopectin, the predominant polysaccharide of starch which differs from glycogen in being less branched and more phosphorylated [222]. Glycogen also includes small amounts of covalent phosphate whose function is poorly understood, at a frequency of 1/ 650 - 1500 glucose residues [15, 216].

As described in the introduction, Lafora disease is a progressive myoclonus epilepsy with teenage-onset and usually death within ten years [122]. The hallmark of Lafora disease is the accumulation of Lafora bodies, deposits that consist of poorly branched glycogen. Increased phosphorylation of glycogen-like polyglucosan in Lafora bodies has been reported [223]. In a mouse model of the disease (*Epm2a-/-* mouse) the gene encoding laforin is disrupted. Increased glycogen phosphorylation has been observed in this mouse line, which led to a derangement of glycogen structure with age, accompanied also by reduced branching and solubility [15, 216]. Glycogen purified from the muscle of 10-12 month old *Epm2a* -/- mice is more phosphorylated and less branched than normal mouse muscle glycogen from the wild type (WT) mice. Stbd1 showed increased co-sedimentation with glycogen purified from *Epm2a* -/- mouse muscle

(Figure 15B). Furthermore, low speed centrifugation (10,000 g) of muscle extracts from *Epm2a* -/- mice essentially separates the normal and aberrant glycogen species into supernatant (low speed supernatant, LSS) and pellet (low speed pellet, LSP). The glycogen purified from LSS is relatively normal, while the poorly branched, more phosphorylated glycogen is in the LSP [15]. We observed that Stbd1 bound more tightly to glycogen purified from the LSP than that form LSS fraction (Figure 15B). Therefore, we conclude that purified Stbd1 can bind glycogen and interacts preferentially with less branched and/or more phosphorylated polysaccharides.

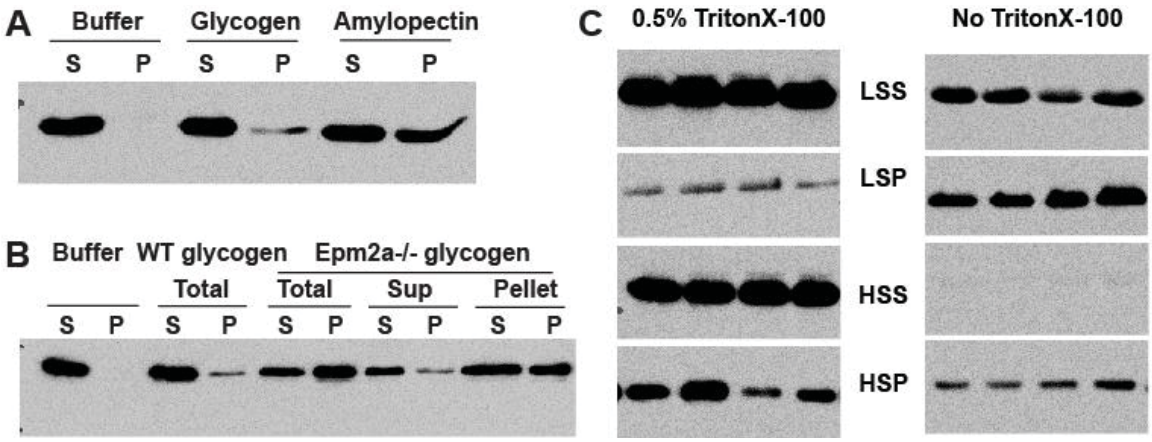


Figure 15. Association of Stbd1 protein with glycogen.

(A) Recombinant human Stbd1 protein binding to glycogen and amylopectin. Rabbit liver glycogen and potato amylopectin (Sigma) were incubated with recombinant hStbd1, centrifuged and analyzed by immnonoblotting using anti-hStbd1 antibodies (anti-GENX-3414). S, supernatant after centrifugation at 100,000 g; P, pellet after centrifugation at 100,000g. The negative control contained Stbd1 but without polysaccharide (Buffer). (B) Recombinant human Stbd1 protein binding to glycogen purified from wild type and laforin knockout mice. Total glycogen was purified from mouse skeletal muscle of wild type (WT) or laforin knockout (*Epm2a -/-*) mice. Muscle extracts from laforin knockout mice was additionally fractionated by low speed centrifugation (10,000 g) to separate supernatant and pellet. Glycogen was purified from this supernatant (Sup) and pellet (Pellet). Purified mouse glycogen (WT total, *Epm2a -/-* total, *Epm2a -/-* Sup and *Epm2a -/-* Pellet) was incubated with recombinant hStbd1, centrifuged at 100,000 g and the supernatant (S) and pellet (P) fractions analyzed by immunoblotting. Stbd1 without added glycogen (Buffer) was the negative control. (C) Fractionation of Stbd1 by centrifugation of wild type mouse muscle extracts in the presence or absence of Triton X-100. Fractions were analyzed by immunoblotting using anti-mStbd1 antibodies, as described under experimental procedures. LSS, supernatant after centrifugation at 8,000 g centrifugation; LSP, pellet after centrifugation at 8,000 g; HSS, supernatant after centrifugation at 100,000 g; HSP, pellet after centrifugation at 100,000 g. The panels shown were cropped from a single autoradiogram of a single membrane.

1.2 Stbd1 interacts with glycogen in muscle extract

Glycogen in muscle extracts can be fractionated by centrifugation, typically by first removing gross cellular debris, including contractile proteins and associated structures, with a low speed centrifugation (10,000 g) followed by ultracentrifugation (100,000 g) of the supernatant to sediment the glycogen particles. In our experiments, about half of the Stbd1 was recovered in the LSP, possibly because of the association with membranous components since by adding a non-ionic detergent Triton X-100 in the homogenization buffer, a greater portion of Stbd1 was present in the LSS fraction (Figure 15C). After high speed centrifugation of the low speed supernatant in the absence of Triton X-100, almost all of the soluble Stbd1 was recovered in the high speed pellet (HSP) fraction (Figure 15C). When the detergent was present, a substantial amount of Stbd1 was still recovered in the HSP although some was left in the high speed supernatant (HSS). Hence, the results indicate that a significant fraction of the Stbd1 in a muscle extract cofractionates with glycogen because of direct binding to the polysaccharide.

1.3 Endogenous Stbd1 distribution in mouse tissues

The commercially available anti-GENX3414 antibody against human Stbd1 (hStbd1) is not able to recognize mouse Stbd1 (mStbd1), probably because the protein sequences of mouse and rat Stbd1 have only 60% identity with that of the human protein. We raised antibodies against recombinant GST-fusion mStbd1 and affinity purified it by binding to immobilized His-tagged mStbd1. Using these antibodies, a single protein species of ~36 kDa was identified in extracts of most mouse tissues tested (Figure 16), including liver, skeletal muscle, adipose tissue, lung, heart and spleen. Liver and muscle, the two major glycogen depositories (Table 2), showed particularly prominent signals of Stbd1 by immunoblotting, consistent with the earlier report [161]. Overexposure of the autoradiogram revealed trace signals in brain, kidney and pancreas.

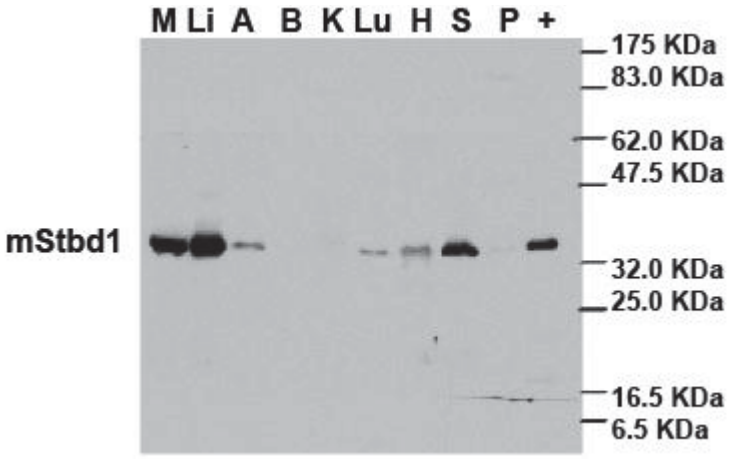


Figure 16. Endogenous distribution of Stbd1 protein in mouse tissues. Stbd1 protein levels in mouse tissues as judged by immunoblotting with anti-mStbd1 antibodies. M, skeletal muscle; Li, liver; A, adipose; B, brain; K, kidney; Lu, lung; H, heart; S, spleen; P, pancreas; +, extract of COS M9 cells over-expressing mStbd1 (running larger because of C-terminal HA-tag). Loadings were 10 µg total protein.

Tissue	Glycogen (µmol glucose / g tissue)
Skeletal muscle	**17.3**
Liver	**384.6**
Adipose tissue	**0.2**
Brain	**2.3**
Kidney	**0.3**
Lung	**1.8**
Heart	**1.9**
Spleen	**1.4**
Pancreas	**0.1**

Irimia, J.M. unpublished data.

Table 2. Glycogen content in different mouse tissues.

1.4 Correlation of Stbd1 with glycogen levels in different genetically modified mice

Muscle glycogen synthase knockout (MGSKO) mice [95] are disrupted in the *Gys1* gene which encodes the isoform of GS expressed in skeletal muscle and most tissues that produce glycogen except liver. Glycogen is undetectable in cardiac and skeletal muscle from survived knockout mice. Immunoblotting of muscle extracts from MGSKO mice had virtually imperceptible Stbd1 protein compared with controls (Figure 17A). Muscle glycogen can also be acutely reduced by 90% of the starting level when mice were exercised to exhaustion on a treadmill [218]. Stbd1 protein levels were not changed by this exercise regimen as judged by immunoblotting of muscle extracts from wild type mice (Figure 17B).

Liver glycogen synthase knockout (LGSKO) mice [93] have a disruption of the *Gys2* gene in the liver, which encodes the liver isoform GS. The liver glycogen level in the fed state of the knockout mice is decreased to 5% of the wild type level. The Stbd1 protein level was reduced by 40% in extracts of livers from fed LGSKO mice compared to control mice (Figure 17C). In addition, the reduction of Stbd1 was seen by immunohistochemical staining of liver sections, in which a visible difference was observed between LGSKO and control (Figure 17D). After overnight fasting, approximately 95% of liver glycogen is degraded in wild type mice, giving a glycogen content comparable to that of fed LGSKO liver [93]. Stbd1 protein levels did not show significant reduction in either WT or LGSKO mice in the fasted state (Figure 17C).

From the data with MGSKO and LGSKO mice, we observed that absence or reduction of glycogen over the long term correlates with decreased Stbd1 protein, whereas short term decreases in either liver or muscle glycogen in wild type mice have no significant effects. The results with the genetically modified mouse lines argue for a genetic link between glycogen and Stbd1. In addition, a strong trend

to elevate Stbd1 mRNA level was observed in the muscle of non-exercised MGSKO mice compared to control WT mice (p=0.0587), while the mRNA level of Stbd1 in the fed LGSKO mouse liver was not significantly changed (Figure 18). Summarily, in the muscle of MGSKO mice where no baseline glycogen content could stabilize Stbd1 protein, the protein is absent. The permanent depletion of glycogen may also serve as a feedback signal to Stbd1 expression, which leads to a tendency of increasing mRNA level. In the LGSKO mice, with 95% reduction of liver glycogen, the Stbd1 protein level was only lowered by 40% without significant change in mRNA level. These observations indicate that Stbd1 protein could have relatively long half-life and can be stabilized by glycogen.

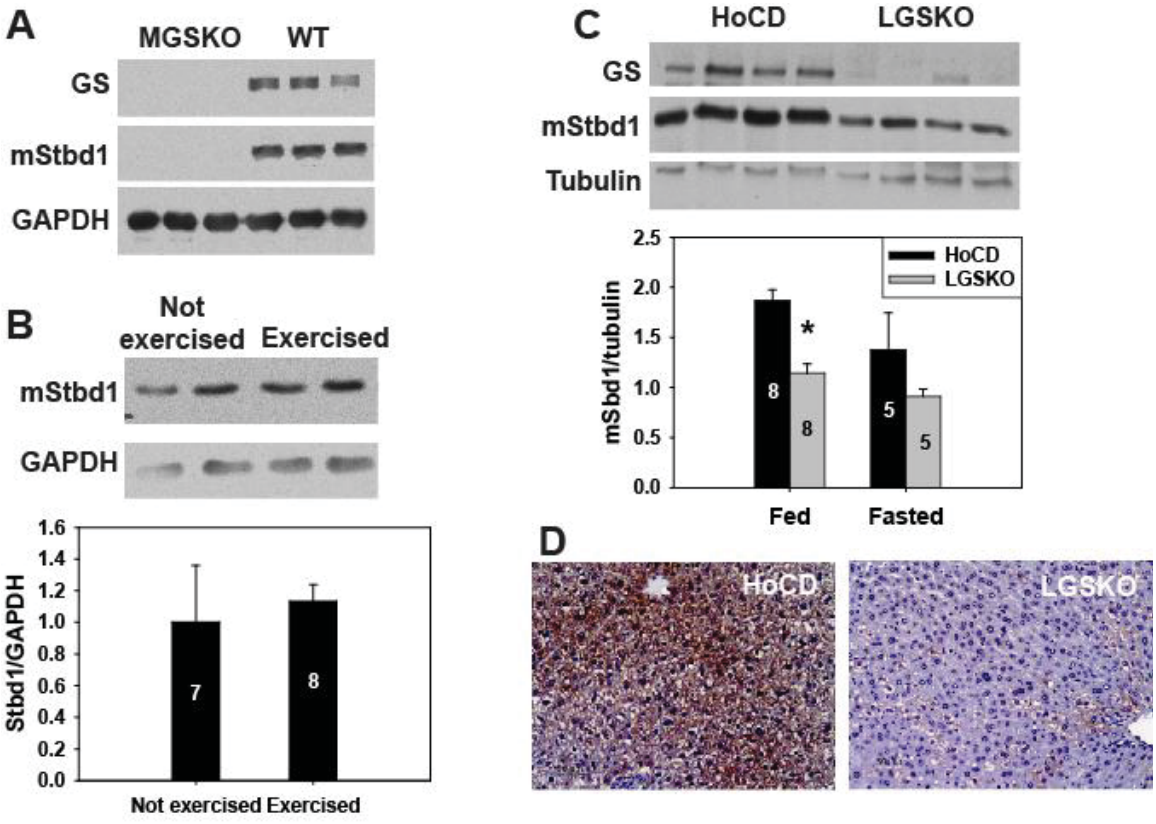


Figure 17. Correlation of Stbd1 protein level with genetic modified glycogen level in mice.

(A) Stbd1 protein level in skeletal muscle from muscle glycogen synthase knockout (MGSKO) mice, with the Gys1 gene disrupted, and control WT

littermates. GS, glycogen synthase; GAPDH, glyceraldehyde-3-phosphate dehydrogenase, the loading control. (B) Stbd1 levels in exercised versus non-exercised wild type mice. A representative Western blot is shown as well as quantitation in the lower bar graph. Numbers within the bars indicate the number of mice analyzed. (C) Representative immunoblots for Stbd1 in livers of fed floxed conditional (HoCD) control mice and liver glycogen synthase knockout (LGSKO) mice, with Gys2 gene disrupted. Quantitation of Stbd1 levels in control or LGSKO mice under fed or fasted conditions is shown in the bar graph below. * $p<0.05$ with respect to HoCD. (D) Stbd1 visualized in liver sections of fed LGSKO and control mice (HoCD) by immunohistochemical staining.

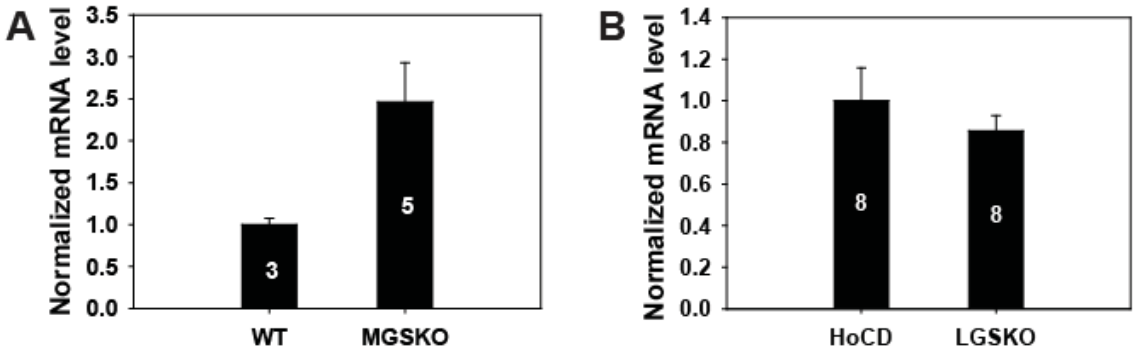


Figure 18. mRNA level of Stbd1 in liver and muscle.

(A) mRNA level in skeletal muscle of non-exercised MGSKO mice and control WT littermates ($p=0.0587$ with respect to WT). (B) mRNA level in liver of fed LGSKO mice and control WT littermates ($p=0.4202$ with respect to HoCD). mRNA content was measured by real time PCR and normalized by reference gene 18 S. Numbers within the bars indicate the number of mice analyzed.

2. Structure-function analysis of Stbd1

Since Stbd1 has been observed binding to polysaccharides *in vitro* and *in vivo* with a preference for poorly branched polysaccharides, and correlating with genetic depletion of glycogen in mouse models, we focused on the structure-function analysis of Stbd1 for further study.

2.1 Sequence analysis

The architecture of Stbd1 indicates an N-terminal hydrophobic segment of 24 amino acids, a putative leucine-zipper domain with three heptad repeats, and a CBM20 domain at the C-terminus (Figure 14) that is conserved from prokaryotes to higher eukaryotes. Human Stbd1 (hStbd1) is 358 residues long with 20 and 25 more amino acids than the mouse (mStbd1) and rat proteins respectively. Mouse and rat Stbd1 share 64-65% identity to the human protein and 89% to each other. To date, orthologs of Stbd1 or predicted Stbd1-like proteins have been identified from zebrafish, chicken to a number of mammals. The number of entries is probably still expanding since when we first analyzed databases three years ago, Stbd1 only showed up in mammals from bioinformatic analysis. The hydrophobic N-terminus and C-terminal CBM20 domains of Stbd1 show a very high degree of conservation (Figure 19). The yeast two hybrid screen using Stbd1 lacking the first 171 residues as bait, picked four Stbd1 clones of different lengths from a human skeletal muscle cDNA library (Figure 20). One clone was the same as the bait. The longest clone was only missing the first six amino acids. The two shortest clones were the same, containing a little more than the CBM20 domain. Therefore, it is likely that Stbd1 can oligomerize by interactions between the CBM20 domains. The putative leucine-zipper domain seems not related to this oligomerization.

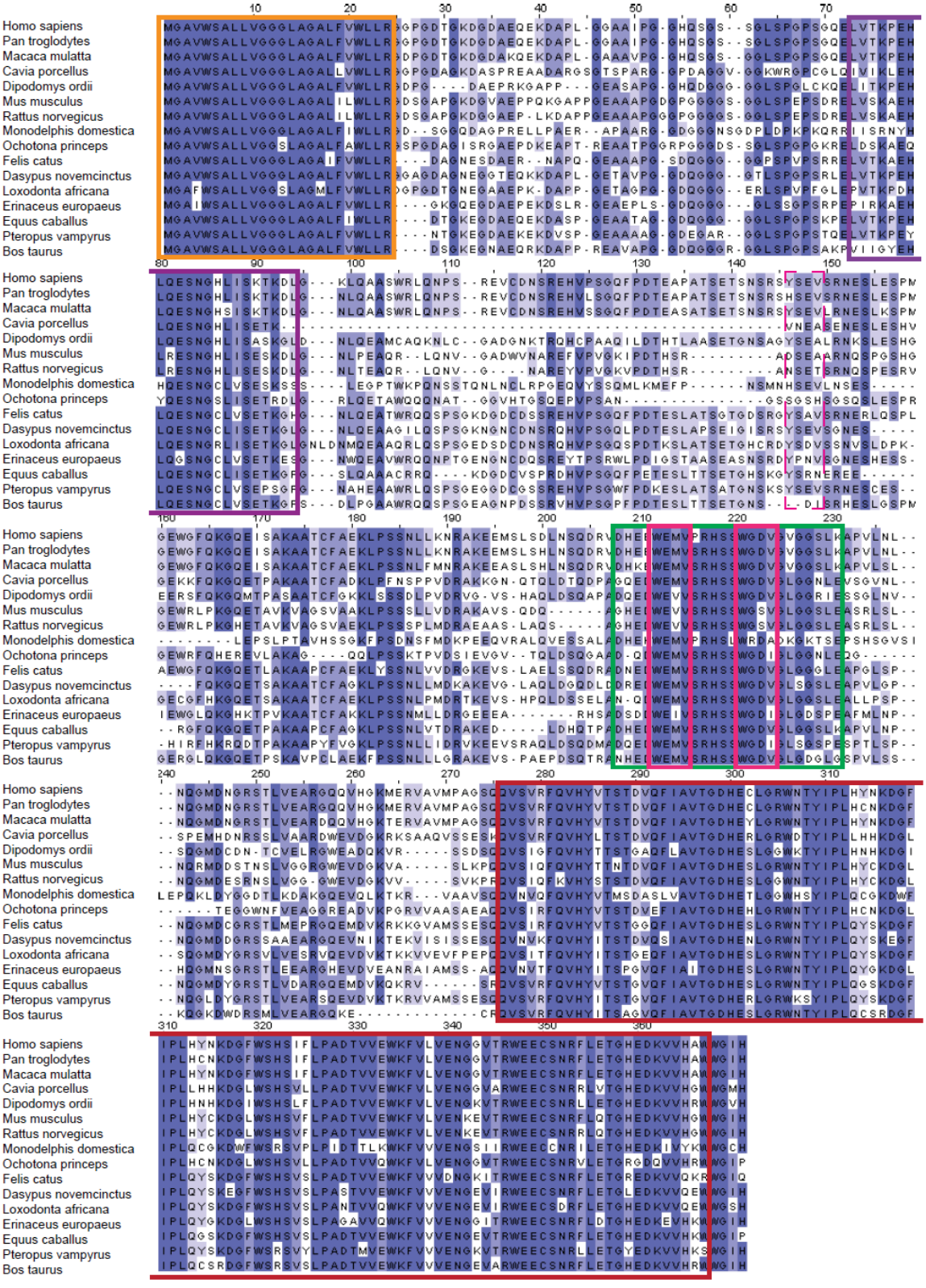

Homo sapiens
Pan troglodytes
Macaca mulatta
Cavia porcellus
Dipodomys ordii
Mus musculus
Rattus norvegicus
Monodelphis domestica
Ochotona princeps
Felis catus
Dasypus novemcinctus
Loxodonta africana
Erinaceus europaeus
Equus caballus
Pteropus vampyrus
Bos taurus

Figure 19. Sequence alignment of mammalian Stbd1.

Sequences were aligned with the Clustal algorithm, with darker shading denoting greater degrees of conservation among species. The orange box indicates the highly conserved hydrophobic N-terminal twenty four residues. The purple box encloses the putative leucine zipper motif. The red box encloses the conserved CBM20 domain. The green box encloses the region includes two potential Atg8 familiy interacting motif (AIM) /LC3 interacting region (LIR) motif. The pink box encloses tow potential AIMs. The pink box with dashed lines encloses a putative AIMs that is excluded due to low sequence conservation.

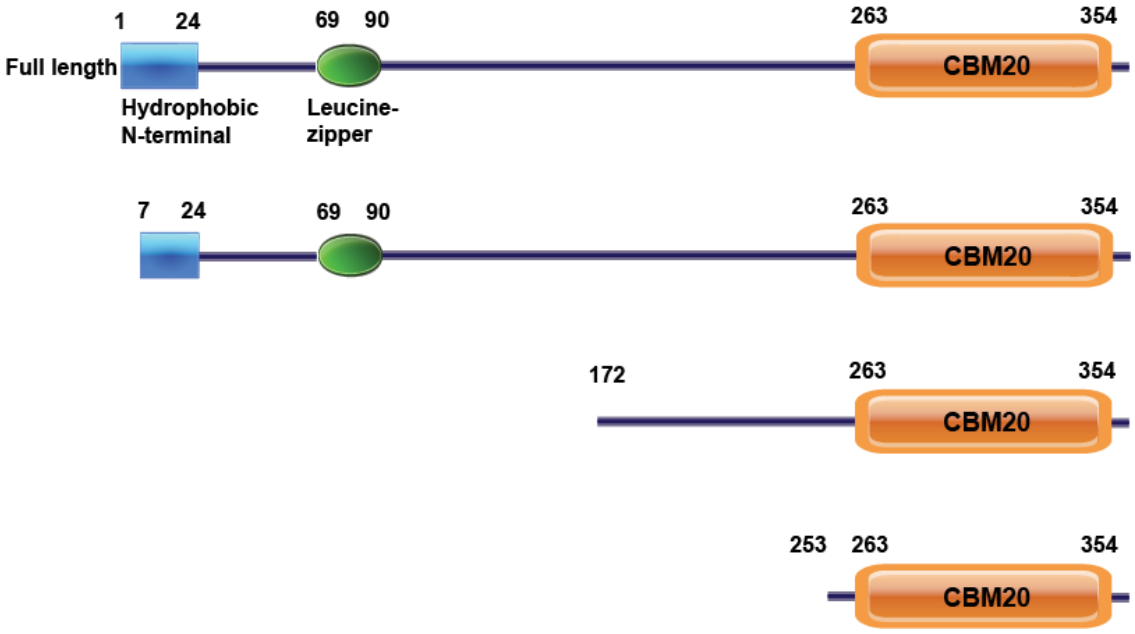


Figure 20. Stbd1 clones of different lengths identified by C-terminal half of Stbd1 in yeast two hybrid screen.

The longest clone is only missing the first six amino acids. One clone without the first 171 amino acids is the same as the bait. The two shortest clones are the same, containing 10 more amino acids than the CBM20 domain.

2.2 Detection of endogenous Stbd1 in cell lines

In several cultured cell lines including mouse C2C12, Hepa1c1c, embryonic fibroblasts (MEFs) and FL83B cells, as well as rat H4IIEC3 and Rat1 cells, we detected endogenous Stbd1 by immunoblotting with the anti-mStbd1 (Figure 21A) but not in human HepG2 or simian COS M9 cells, which again may be explained by relatively low sequence identities of Stbd1 in rodents and primates.

In either the mouse liver cell line FL83B [224] or a Rat1 fibroblast cell line, Rat1Neo5 [217], endogenous Stbd1 was detected by immunofluorescence, to appear predominantly concentrated at perinuclear structures, with diameter up to ~0.5 μm (Figure 21B). These structures partly coincide with the presence of an ER marker, residues HDEL, detected by anti-HDEL (Figure 22). Endogenous Stbd1 also exists within areas with the densest microtubular network (Figure 22).

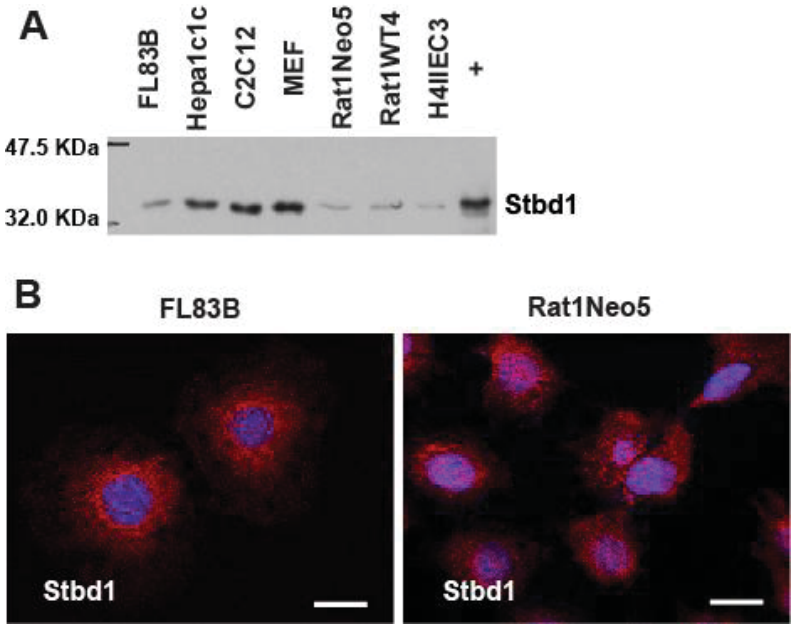


Figure 21. Endogenous Stbd1 in mouse and rat cells.

(A) Presence of Stbd1 protein in mouse and rat cell lines, as described in the text. +, COS M9 cells over-expressing mStbd1. (B) FL83B (left panel) and RatNeo5 (right panel) cells were immunostained with anti-mStbd1 antibodies (red) to detect endogenous Stbd1.

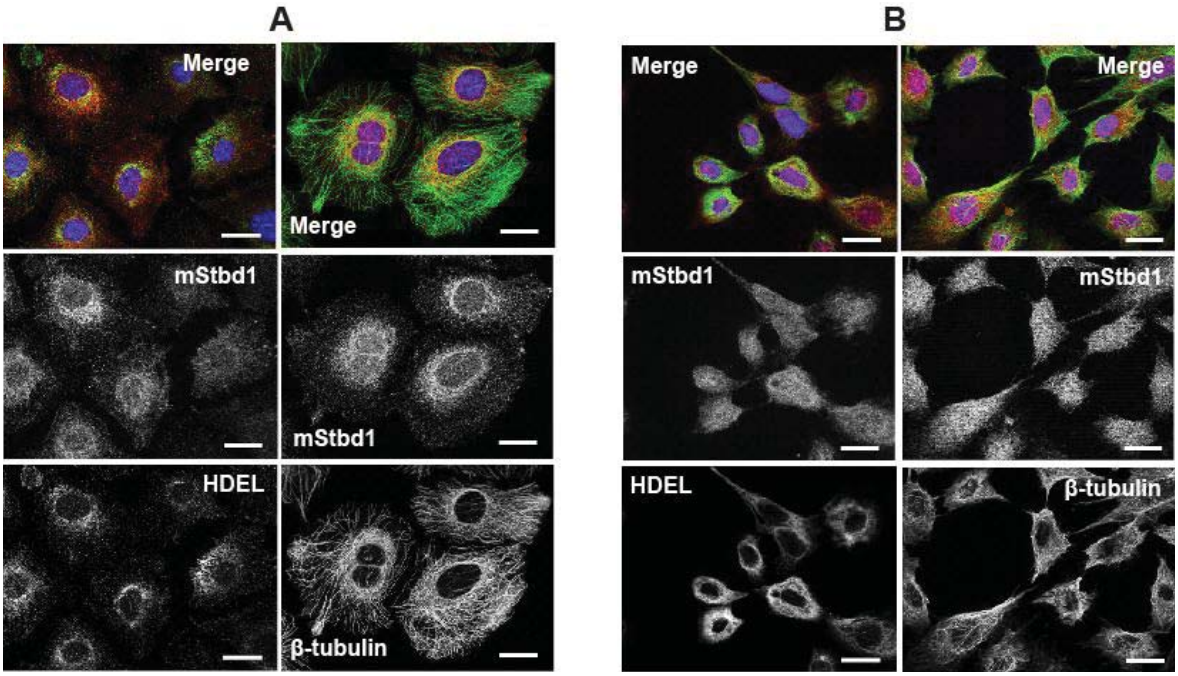


Figure 22. Subcellular localization of endogenous Stbd1 with respect to organelle markers in FL83B and Rat1Neo5 cells.

(A) FL83B cells were immunostained with antibodies directed towards mStbd1 (middle panels) and antibodies towards HDEL, as an endoplasmic reticulum marker (lower left panel), or against β-tubulin as a microtubule marker (lower right panel). The upper panels show merges of the images, with Stbd1 (red) and the corresponding organelle marker (green), with nuclei stained with Hoechst (blue). (B) Rat1Neo5 cells were immunostained with antibodies directed towards mStbd1 (middle panels) and antibodies towards, HDEL, as an endoplasmic reticulum marker (lower left panel), or against β-tubulin as a microtubule marker (lower right panel). The upper panels show merges of the images, with Stbd1 (red) and the corresponding organelle marker (green), with nuclei stained with Hoechst (blue). The scale bars are 20 μm.

2.3 Structure and function analysis by study of hStbd1-HA and mutants

When the HA-tagged full length human Stbd1 (hStbd1-HA) was transiently expressed in COS M9 cells, the protein was located in perinuclear structures with diameters up to several μm (Figure 23A, 24). Both syntaxin 6, a trans-Golgi marker and HDEL (His-Asp-Glu-Leu) that labels ER, partially co-localized with these structures (Figure 24). hStbd1-HA did not co-distribute with microtubules as visualized by an antibody against β-tubulin. Stbd1 tended to be present in areas with the densest network of microtubules (Figure 24).

To analyze the role of different domains of Stbd1 in its subcellular distribution, we applied mutational analysis. Removal of the conserved hydrophobic N-terminal 24 amino acids (ΔN24-HA) resulted in a diffuse cytosolic distribution of Stbd1 as well as complete loss of the large perinuclear structures (Figure 23B). This observation indicates that membrane association of Stbd1 is necessary for the formation of the enlarged perinuclear structures. Further N-terminal truncation that deletes the putative leucine-zipper domain as well (ΔN90) gave an identical pattern (Figure 23C). These results are consistent with the hydrophobic region having an important role in sub-cellular localization, most likely by directing the protein to membrane compartments.

At the C-terminus, Stbd1 has a putative carbohydrate binding domain of the CBM20 family that we proved its polysaccharide binding ability. Here we analyzed its contribution to the subcellular localization of the protein. This domain is highly conserved in Stbd1 in mammalian species according to our protein sequence alignment. Interestingly, when we removed the carbohydrate binding domain (ΔC96), the perinuclear localization is also disrupted (Figure 23D). However, compared to that of the N-terminal deletion, the shortened protein was distributed in a different way with a more reticular appearance perhaps because the loss of the CBM20 domain disabled cargo anchoring and/or Stbd1 oligomerization. Thus, the sub-cellular distribution of C-terminal truncation

of Stbd1 may result from nonspecific insertion of Stbd1 into membranes via the hydrophobic N-terminus.

Our experiments have not indicated any function for the putative leucine-zipper motif in Stbd1. The yeast two hybrid screen using as bait Stbd1 lacking the leucine-zipper region identified Stbd1 as a target, implying that the leucine-zipper region is not necessary for Stbd1 oligomerization. Furthermore, an N-terminal truncation removing both the hydrophobic segment and the leucine-zipper displayed the same diffuse subcellular distribution as removal of the hydrophobic segment alone.

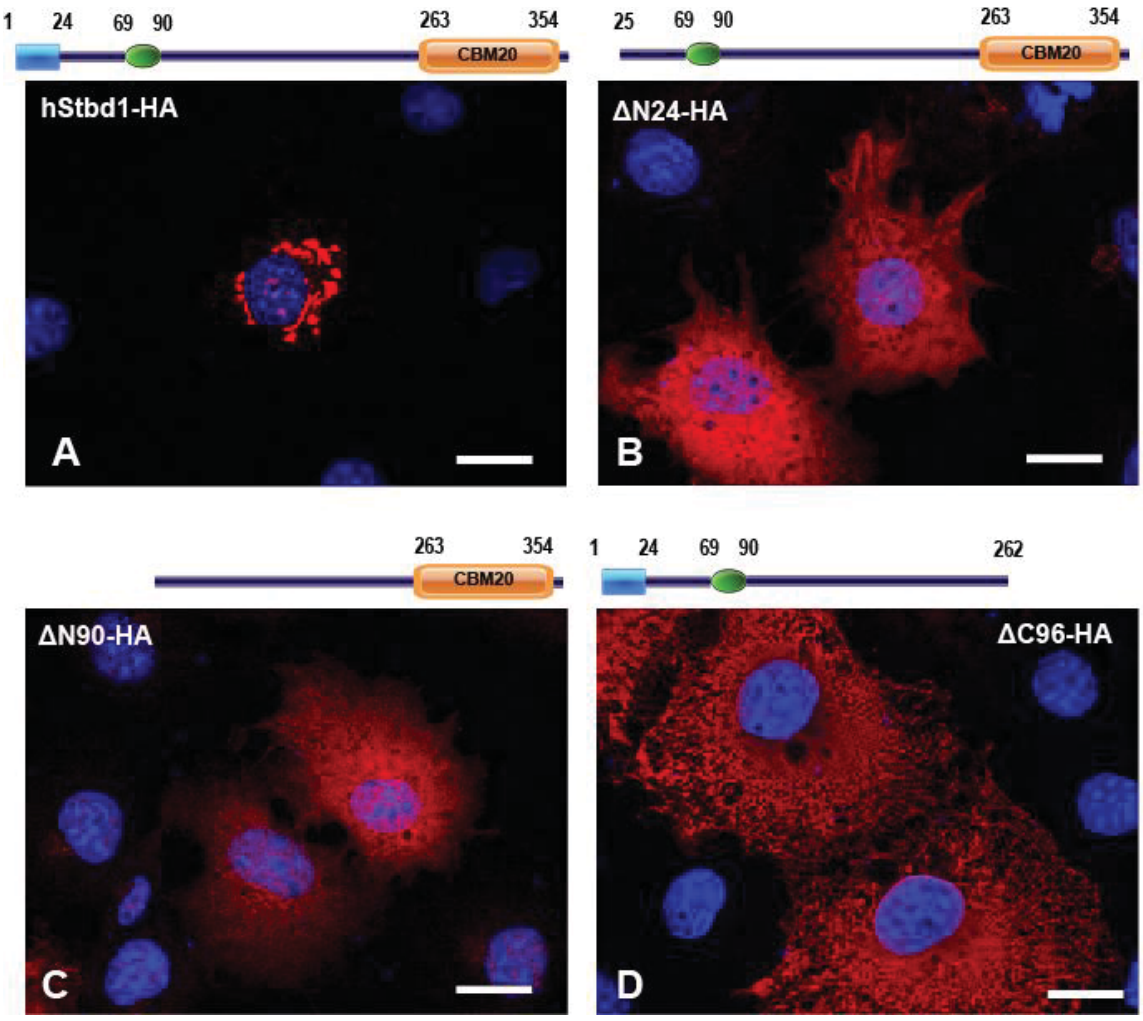


Figure 23. Subcellular localization of Stbd1.

Mutated hStbd1 with a C-terminal HA-tag was expressed in COS M9 cells and immunostained with anti-HA antibodies (red). (A) Full length human Stbd1. (B) Deletion of the N-terminal hydrophobic segment (ΔN24-HA). (C) Deletion of the N-terminal hydrophobic segment and the putative leucine zipper (ΔN90-HA). (D) Deletion of the C-terminal carbohydrate binding domain (ΔC96-HA. Nuclei were stained with Hoechst (blue). The scale bar is 20 μm.

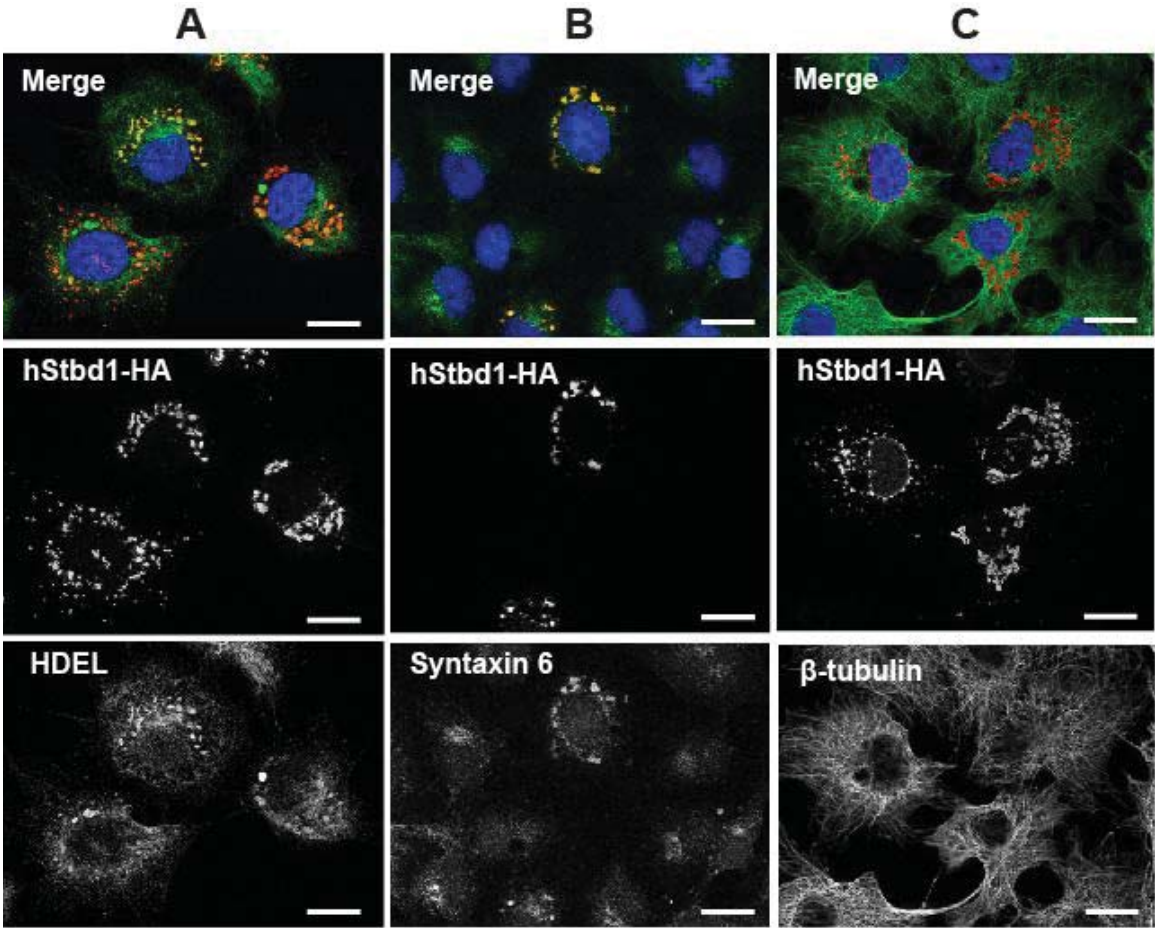


Figure 24. Subcellular localization of Stbd1 over-expressed in COS M9 cells with respect to organelle markers.

Cells overexpressing full-length hStbd1 were immunostained with anti-HA antibodies (middle panels) and antibodies towards HDEL as an endoplasmic reticulum marker (A, lower panel), towards syntaxin 6 as a Golgi marker (B, lower panel) or against β-tubulin as a microtubule marker (C, lower panel). The upper panels show merges of the images, with Stbd1 (red) and the corresponding organelle marker (green), with nuclei stained with Hoechst (blue). The scale bars are 20 μm.

2.4 Glycogen association of Stbd1 in cells and study of glycogen binding site on Stbd1

Though various lines of evidence indicate a correlation between Stbd1 and glycogen, it is important to directly demonstrate this association in a cellular environment. It is difficult to find a good cell line to further analyze co-localization of glycogen and Stbd1 by light microscopy since most cultured cells do not accumulate substantial amounts of glycogen that can be visualized by light microscopy. Therefore, we utilized Rat1WT4 cells which stably overexpress wild type rabbit muscle glycogen synthase [217]. The Rat1Neo5 cells described earlier were control cells made in the same study by transfection with empty vector. When grown under normal conditions of 25 mM glucose, glycogen could be detected in the Rat1WT4 cells by periodic acid/Schiff (PAS) staining, visualized as very large deposits, usually no more than a few per cell and proximal to the nucleus (Figure 25A). These large glycogen deposits were observed to co-localize with Stbd1 (Figure 25B) using the modified method that combines PAS and immunefluorescent staining on the same fixed sample [220]. When we incubated the Rat1WT4 cells in glucose free medium for 24 hours after initial growth in medium with 25 mM glucose, the massive glycogen deposits dissipated into much smaller but more numerous punctate PAS-positive structures that showed substantial co-localization with Stbd1 (Figure 25A and B). Immunofluorescent staining for GS in Rat1WT4 cells gave a clear ring-like appearance in many cells and the enzyme was visualized coincident with the massive glycogen deposits in an enclosure way (Figure 25C). In some cases, the GS defined a loop within which neither Stbd1 nor PAS-positive material was detected, as though the glycogen had been evacuated at some stage along with the Stbd1, leaving the GS behind.

In order to study the glycogen binding site of Stbd1, point mutants of Stbd1 were also constructed, in which a highly conserved Trp (W293 in human) in the CBM20 domain was replaced by either Gly or Leu with the objective of disabling

glycogen binding. W293 was selected for two reasons. Firstly, three conserved tryptophans within CBM20 have been suggested as critical amino acids for polysaccharide binding and W293 aligns with the first pivotal tryptophan of binding site 1 which has been suggested as the primary recognition site of starch in GA SBD from *A. niger,* the best studied CBM20 structure. Secondly, unlike most proteins containing CBM20s which are microbial origin, Stbd1 and laforin are two mammalian proteins that contain CBM20. A mutation at the corresponding conserved W32 in the laforin CBM20 domain, W32G, had been found in a Lafora patient and shown not to bind to glycogen [225, 226]. The W293G and W293L mutants appeared perinuclear without causing the formation of the larger and more well-defined structures seen with the wild type protein (Figure 26).

From our experience, PAS staining does not readily detect the low level of glycogen normally present in Rat1Neo5 or COS M9 cells, most likely due to its dispersed distribution throughout the cells. However, in COS cells overexpressing Stbd1 showing the perinuclear structures described above, punctate glycogen deposits were highlighted by PAS staining (Figure 27A). The total glycogen levels in COS cells measured biochemically was unchanged by over-expression of Stbd1 (Figure 27B). When medium lacking glucose replaced the normal medium, the PAS-positive structures dissipated consistent with the glycogen being degraded as an energy source (Figure 27A). Meanwhile, the combination of immunofluorescent staining for Stbd1 and PAS staining for glycogen demonstrated their co-localization in these perinuclear structures (Figure 28). In contrast, although the W293G and W293L mutants were readily detected in a perinuclear location, as in Figure 26, no signal for PAS staining was detected with mutant proteins (Figure 28), suggesting that the mutant Stbd1 was unable to co-localize with glycogen. W293 aligns to one of the pivotal Trp residue in the microbial starch binding site I and W32 that is required for glycogen binding in mammalian protein laforin. Hence, the slightly different appearance between the hStbd1-HA and W293G-HA or W293L-HA positive

structures is probably due to the absence of glycogen. Therefore, we infer that the glycogen is concentrated by Stbd1 in these perinuclear structures.

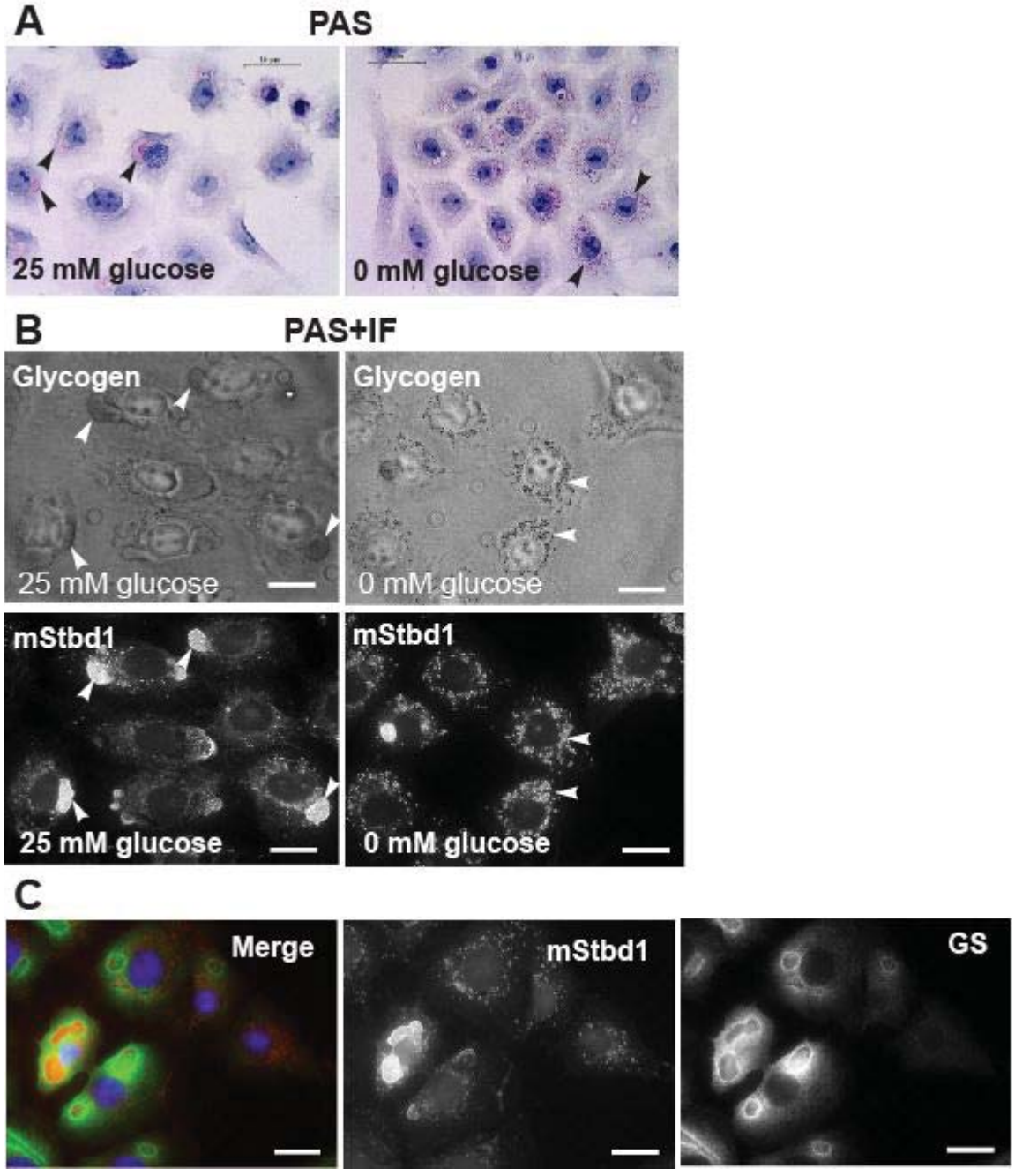


Figure 25. Co-localization of Stbd1 with glycogen in Rat1 cells.

(A) Rat1WT4 cells grown under normal glucose (25 mM; left panel) or starved (0 glucose; right panel) conditions for 24h were stained with periodic acid/Schiff reagent (PAS) to visualize glycogen (pink) by light microscopy. Arrowheads

indicate examples of glycogen staining. (B) Co-localization of endogenous Stbd1 protein and glycogen. Rat1WT4 cells were grown as in (A), subjected to PAS staining to visualize glycogen (upper panels) followed by immunofluorescent staining to detect Stbd1 (lower panels). Arrowheads indicate examples of the co-localization of glycogen and Stbd1. (C) Rat1WT4 cells were immunostained for Stbd1 (middle panel) and glycogen synthase (GS) (right panel). The merged image shows Stbd1 (red) and GS (green) with nuclei stained with Hoechst (blue).

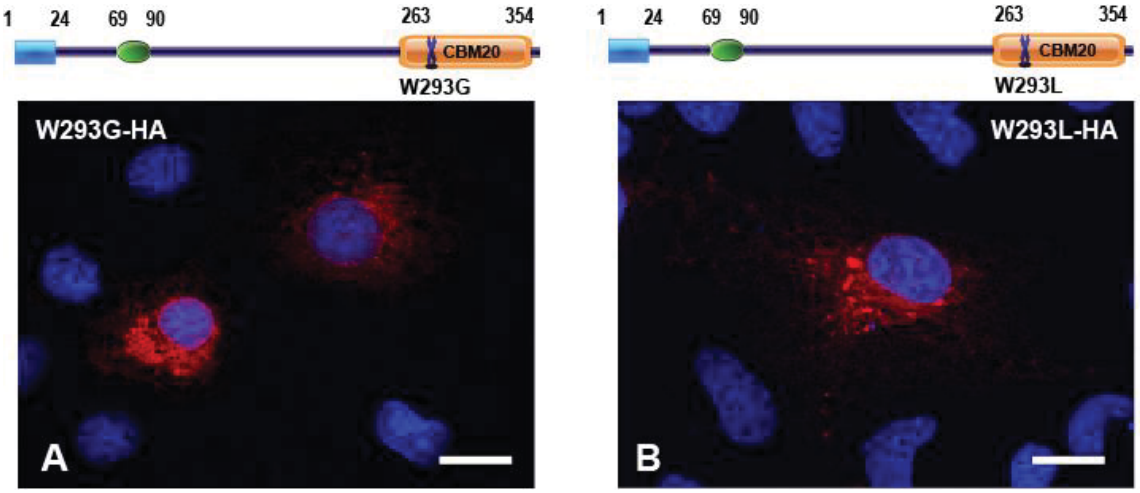


Figure 26. Point mutational analysis of Stbd1 expressed in COS M9 cells. (A) Mutation of conserved Trp of CBM20 domain to Gly (W293G). (B) Mutation of conserved Trp of CBM20 domain to Leu (W293L). Nuclei were stained with Hoechst (blue). The scale bar is 20 μm.

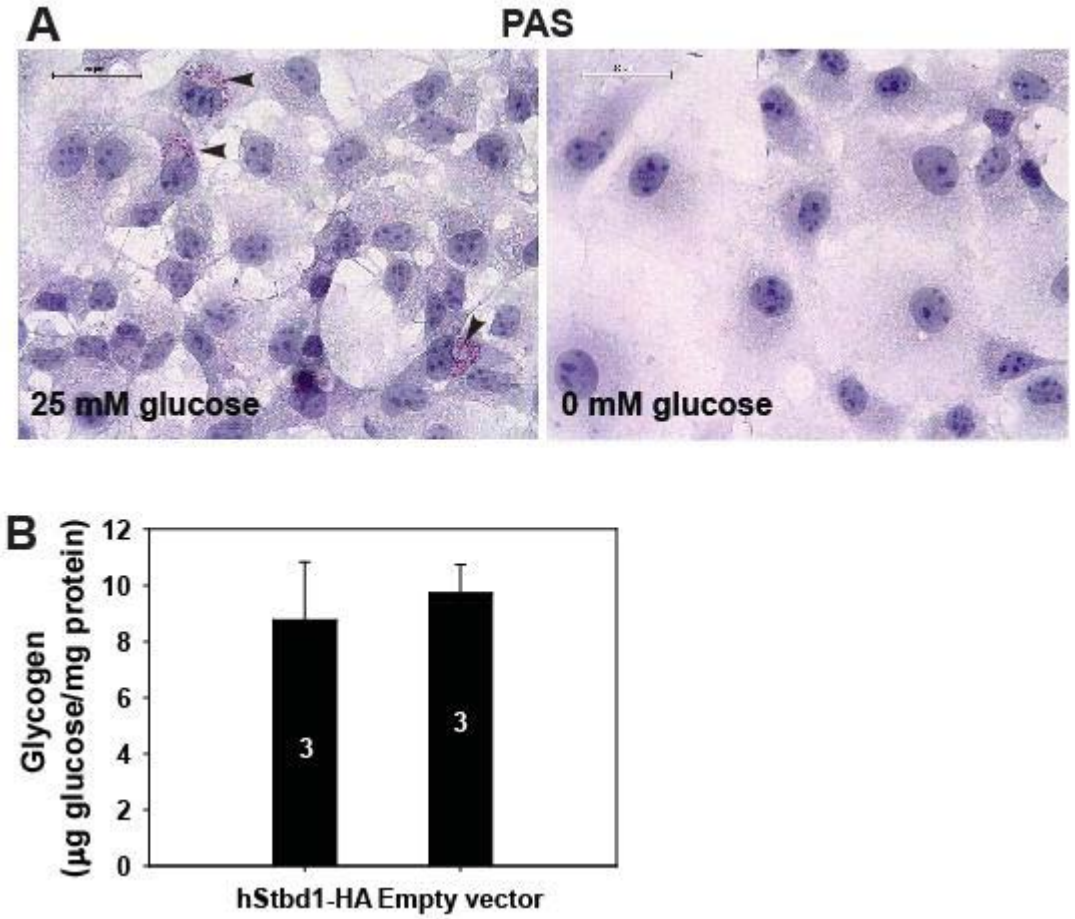


Figure 27. Detection of glycogen in COSM9 cells with or without overexpressing Stbd1.

(A) COS M9 cells overexpressing human Stbd1 (hSTBD1-HA) were incubated in medium with 25 mM glucose (upper panel) or no glucose (lower panel) for 24h and stained with PAS to visualize glycogen by light microscopy. Arrow heads indicate examples of glycogen staining in the fed cells. (B) Glycogen content in COSM9 cells with or without overexpressing hStbd1-HA.

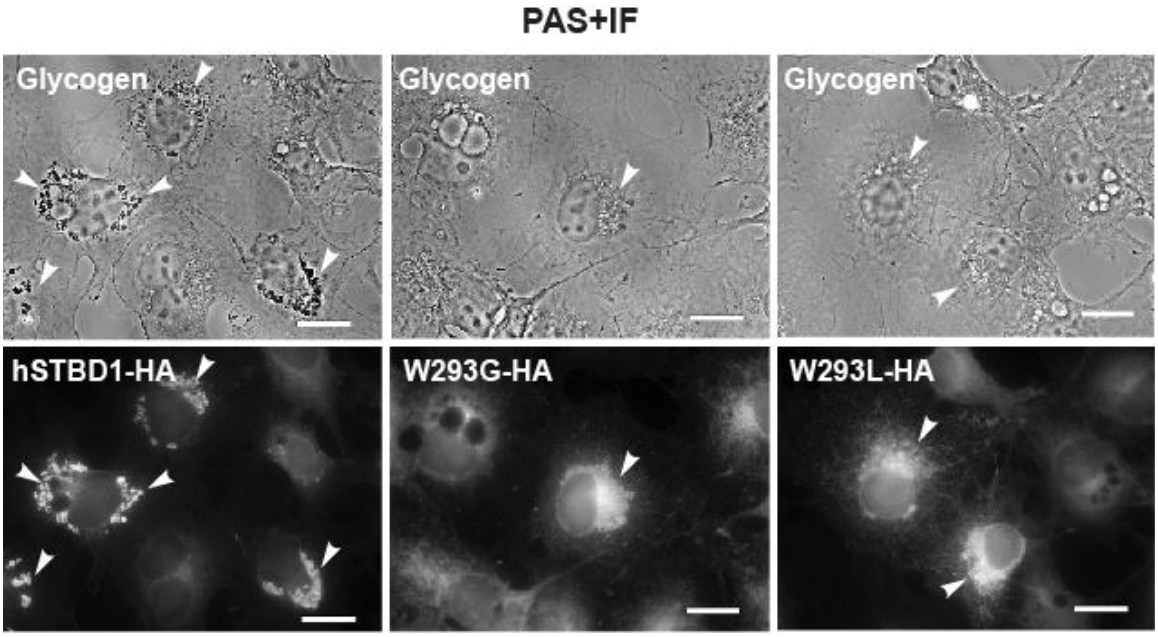


Figure 28. Co-localization of Stbd1 with glycogen in COS cells.

COS M9 cells were over-expressed with full length human Stbd1 (hSTBD1-HA; left panels) or the CBM20 point mutants (W293G-HA; middle panels and W293L-HA; right panels). Cells were subjected to PAS staining to visualize glycogen under bright field (upper panels) followed by immunofluorescent staining with anti-HA antibodies to visualize Stbd1 under fluorescent channel (lower panels). Arrowheads in the lower panels indicate Stbd1 localization and, in the upper panels, the colocalized glycogen staining with full length Stbd1. The scale bars are 20 μm.

3. Stbd1 interacting proteins

As we learned from our study that Stbd1 binds and tracks glycogen in cells, understanding the glycogen trafficking mechanism was our next goal. We used yeast two hybrid screen to identify potential interacting proteins of stbd1. Using Stbd1 lacking the first 171 residues as bait, several interesting targets were picked from a human skeletal muscle cDNA library, including four Stbd1 clones of different lengths as mentioned before, two autophagy related proteins, GABARAP and GABARAPL1, the glycogen phosphatase laforin .and a transcription factor cut-like homeobox 1.

3.1 Interaction of Stbd1 with GABARAPL1 and/or GABARAP

The same yeast two hybrid screen that suggested Stbd1 oligomerization also identified GABARAP and GABARAPL1 as potential interacting proteins. These two small proteins are members of the ATG8 family that are involved in autophagy [179, 213].

3.1.1 Confirmation of interaction

The interaction of Stbd1 with each protein was confirmed by co-immunoprecipitation of the proteins expressed with epitope tags tags in COS cells (Figure 29), whether the pull-down was of the ATG8 protein via an N-terminal Flag tag or of Stbd1 via a C-terminal HA tag. A recent proteomic analysis of human proteins involved in autophagy also reported interaction of Stbd1 with GABARAP and GABARAPL1 [167].

Immunofluorescent staining of co-expressed proteins in COS cells indicated a strong co-localization of hStbd1-HA with Flag-GABARAPL1 in the perinuclear structures described above (Figure 30A). There was some co-localization of Stbd1-HA with Flag-GABARAP but not as consistently or as strictly as with Flag-

GABARAPL1 (Figure 30A). To confirm the interaction of GABARAPL1 with Stbd1, we also analyzed the sub-cellular distribution of endogenous GABARAPL1. In cells transfected with control vector, GABARAPL1 showed a punctate appearance throughout the cytosol (Figure 30B). When hStbd1-HA was overexpressed, endogenous GABARAPL1 was restricted to the perinuclear structures (Figure 30B), similar to what was seen with Flag-GABARAPL1 over-expression (Figure 30A). Endogenous LC3, another member of the Atg8 family, however, gave a weak signal and did not co-localize with overexpressed Stbd1 (Figure 30C) which indicated that the interaction of Stbd1 and GABARAPL1 could be specific.

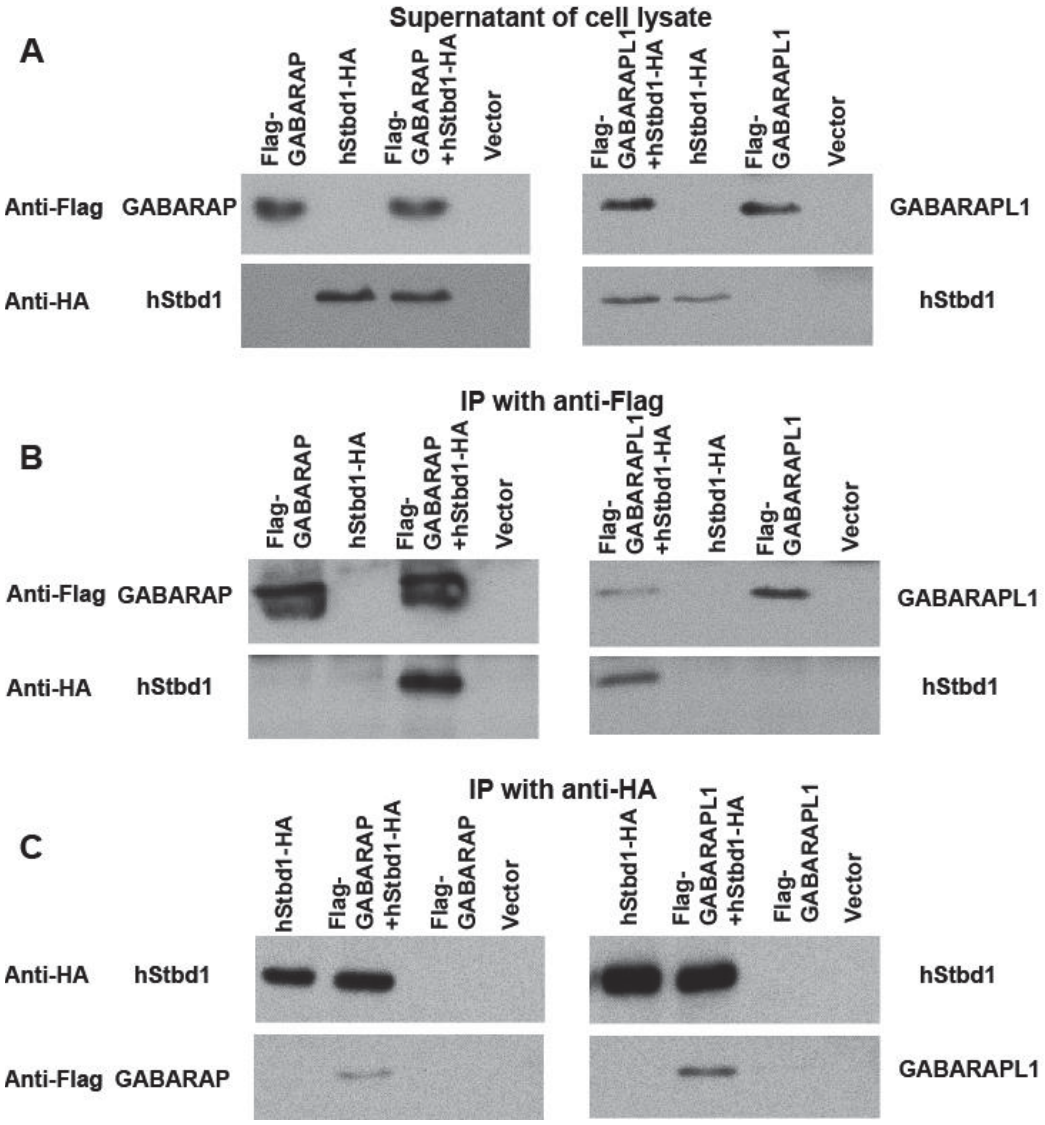


Figure 29. Interaction of Stbd1 with GABARAPL1 and GABARAP.
hStbd1 with a C-terminal HA-tag, N-terminal Flag-tagged GABRAPL1 and N-terminal Flag-tagged GABARAP were expressed alone, or in the indicated combination in COS M9 cells. Control cells were transfected with empty pcDNA3 vector (Vector). (A) Immunoblotting of the cell lysates with the indicated antibody. (B) Immunoprecipitation of GABARAP (left) or GABARAPL1 (right) with Anti-Flag antibodies covalently bound to agarose followed by immunoblotting with the indicated antibody. (C) Immunoprecipitation of Stbd1 with anti-HA antibodies covalently bound to agarose followed by immunoblotting with the indicated antibody.

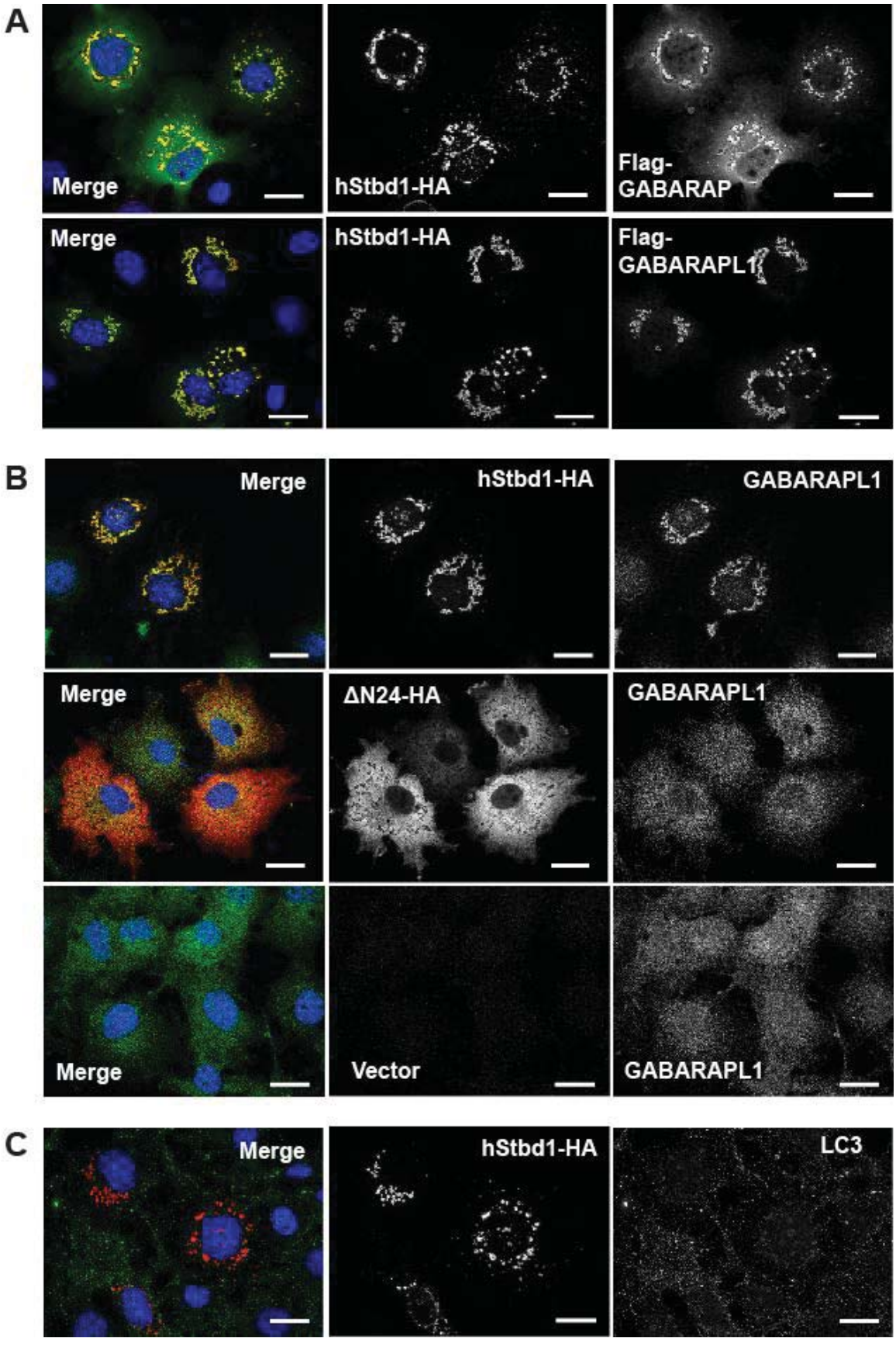

A
Merge
hStbd1-HA
Flag-
GABARAP
Merge
hStbd1-HA
Flag-
GABARAPL1
B
Merge
hStbd1-HA
GABARAPL1
Merge
ΔN24-HA
GABARAPL1
Merge
Vector
GABARAPL1
C
Merge
hStbd1-HA
LC3

Figure 30. Subcellular localization of GABARAPL1, GABARAP and LC3 in relation to Stbd1.

(A) C-terminal HA-tagged hStbd1 was co-expressed in COS M9 cells with N-terminal Flag-tagged GABARAPL1 or GABARAP, as in Figure, and immunostained with anti-HA antibodies (red) or anti-Flag antibodies (green). (B) HA-tagged Stbd1 or N-terminally truncated Stbd1 was expressed in COS M9 cells and immunostained with anti-HA antibodies to detect Stbd1 (red) and anti-GABARAPL1 antibodies to visualize endogenous GABARAPL1 (green). The bottom row shows cells transfected with empty vector (pcDNA3) to reveal the endogenous GABARAPL1 distribution (green). (C) HA-tagged Stbd1 was expressed in COS M9 cells and immunostained with anti-HA antibodies (red) or anti-LC3 antibodies (green) to detect endogenous LC3. Nuclei were stained with Hoechst (blue). The scale bar is 20 μm.

3.1.2 Identification of Atg8 family interacting motifs on Stbd1

Studies have shown that proteins interact with Atg8 family proteins via Atg8 family interacting motif, the AIMs, typically a WxxL-like sequence [213] (Figure 8). To further understand the structural basis of Stbd1 and GABARAPL1 and/or GABARAP interaction, several putative AIMs in Stbd1 were identified, localized throughout the sequence (Figure 31) based on the basic sequence pattern.

Using the truncation mutants we already constructed, we were able to narrow down the possible region of Stbd1 that mediates the association with autophagy related proteins. In the same COSM9 cell system, Flag-GABARAPL1 was co-expressed with Stbd1 mutant ΔN24-HA removing the N-terminus or ΔC96-HA lacking the C-terminus. The immunoprecipitaion results indicated that neither the N-terminal hydrophobic segment nor the C-terminal CBM20 domain of Stbd1 is required for its association with GABARAPL1 since both mutants were able to co-immunorecipate with this Atg8 family member just like the full length protein (Figure 32). Furthermore, an immunofluorescent experiment revealed that the hStbd1 mutant ΔN24-HA abolished the protein accumulation in the perinuclear structures as shown before (Figure 23B) but still co-localized with endogenous GABARAPL1 (Figure 30B), consistent with the yeast two-hybrid screen in which GABARAPL1 was identified using N-terminally truncated Stbd1 as the bait. Moreover, analysis of the Stbd1 sequence for inherent disorder using the predictors of natural disordered regions (PONDR) algorithm [227] suggested that the entire region between the N-terminal hydrophobic segment and the C-terminal CBM20 domain has a high probability of disorder and possibly an increased possibility of being involved in protein-protein interactions interaction. As more and more Atg8 family interacting proteins have been identified and studied, the AIM can be refined as $x_{-3}x_{-2}x_{-1}$W/F/Yx_1x_2L/I/V, where X_1 and/or X_2 and at least one of X_{-3}-X_{-2}-X_{-1} would be acidic residues

At this point, there remained three possible AIMs, SRSYSEV , HEEWEMV (W203 and V206 in human Stbd1) and HSSWGDV (W212 and V215 in human Stbd1) in the middle of Stbd1 protein sequence. The SRSYSEV sequence was excluded because it was poorly conserved in different species (Figure 19). W203, V206 and W212 are invariant in all 16 mammalian Stbd1 protein sequences in our alignment. V215 is also highly conserved in the alignment, with only 4 out of 16 exceptions, but the Val residue was replaced by Ile, another large hydrophobic residue (Figure 19). Therefore, we constructed a deletion Δ198-222 that removed the conserved region containing both putative AIMs of Stbd1, two double mutations, W203A+V206A and W212A+V215A and four single point mutations, W203A, V206A, W212A and V215A, hoping to define the interaction site(s) (Figure 14). All the mutants had HA tag at their C-termini like the full length protein that we had studied. Each mutants was individually overexpressed in COSM9 cells and the familiar perinuclear enlarged vesicle-like structure was observed (Figure 33).

In the immunoprecipitation as well as immunoflurorescent staining experiments, Flag-GABARAPL1 was co-expressed with each mutant. First, the partial deletion, Δ198-222-HA lost the ability to co-immunoprecipitate with Flag-GABARAPL1 (Figure 34) and the following immunofluorescent staining of the two proteins showed no co-localization (Figure 35B) even though Δ198-222-HA appeared in the familiar perinuclear structures. When the two double mutations were applied in these tests, only (W203A+V206A)-HA reproduced what we observed with Δ198-222-HA (Figure 35C). (W212A+V215A)-HA acted like wild type full length Stbd1-HA (Figure 35D). Mutation of either W203 or V206 in Stbd1 is sufficient to eliminate GABABARAL1 association (Figure 36A and B). In contrast, the other single mutants W212A-HA and V215A-HA behaved like wild-type Stbd1, in terms of subcellular co-localization with Flag-GABARAPL1 (Figure 36C and D). The W293G-HA, point mutation in CBM20 domain showed different perinulcear pattern comparing to wild type Stbd1, but kept the co-locolization with Flag-GABARAPL1 (Figure 36E), suggesting the loss of interaction is specific.

Furthermore, the endogenous GABARAPL1 showed similar distribution pattern to Flag-GABARAPL1 overexpression in cells. It was no longer co-localized with Δ198-222 HA and (W203A+V206A)-HA while remained coincident with (W212A+V215A)-HA (Figure 37). These data indicates that WEMV (203-206) acts as AIM on Stbd1 for GABARAPL1 interaction.

3.2 Potential interaction between Stbd1 and laforin

Another interesting protein target identified from the same yeast two hybrid screen is the glycogen phosphatase laforin which also has a CBM20 domain, although in this case at its N-terminus. When co-expressed with Flag-laforin in COSM9 cells, hStbd1-HA co-immunoprecipitated Flag-laforin. However, pull-down of laforin did not give co-immunoprecipitation of hStbd1-HA. Stbd1 oligomerizes via its CBM20 domain and perhaps interacts with laforin via a similar interaction. Then, HA antibodies binding to the C-terminal HA tag of hStbd1 might interfere with its association with laforin. In immunofluorescent staining data, Flag-laforin was not consistently present in the perinuclear structures of hStbd1-HA, possibly because laforin targets cytosolic glycogen with preference to membrane-bound glycogen. More work is needed to understand whether there is any interaction between Stbd1 and laforin, though it would not be unreasonable since Stbd1 binds preferentially to phosphorylated glycogen that is presumably a substrate for laforin.

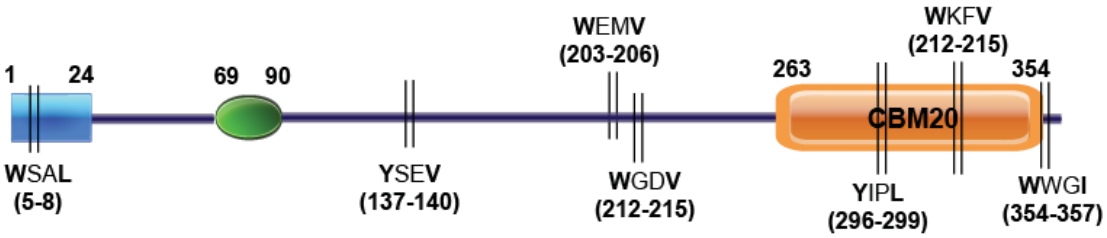


Figure 31. Schematic architecture of human Stbd1 with potential AIMs.

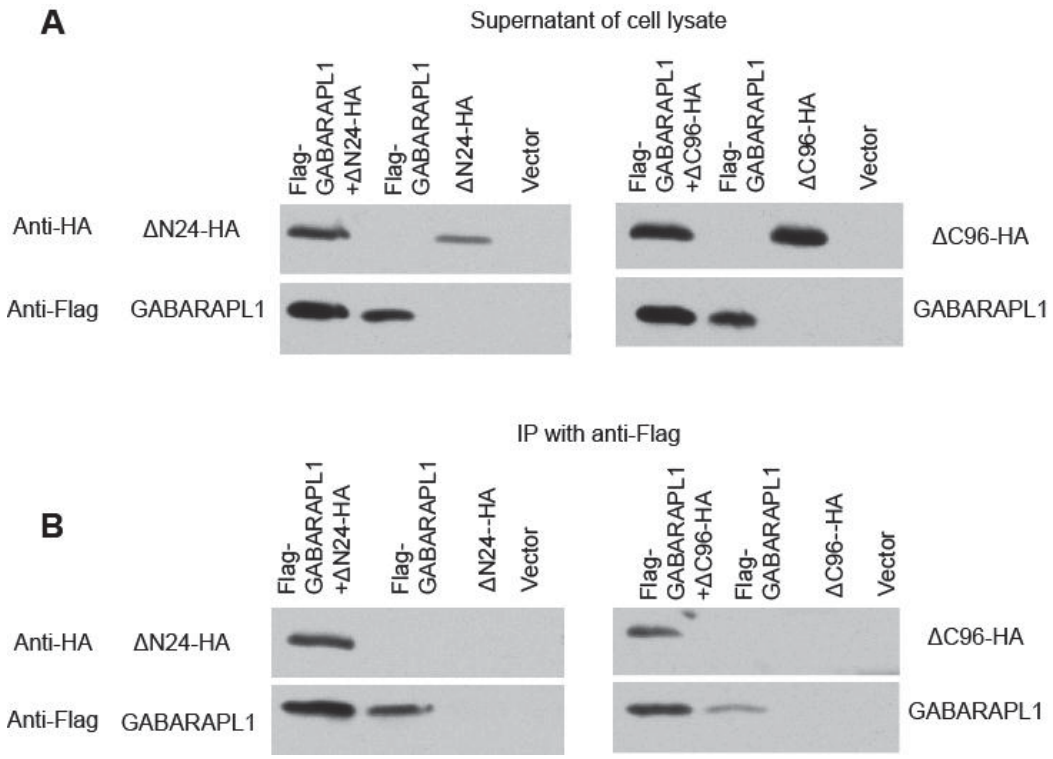


Figure 32. Interaction of GABARAPL1 with truncated mutants of Stbd1. Truncated hStbd1 lacking 1-24 residues or 262-358 residues with a C-terminal HA-tag (ΔN24-HA or ΔC96-HA) and N-terminal Flag-tagged GABRAPL1 were expressed alone, or in the indicated combination in COS M9 cells. Control cells were transfected with empty pcDNA3 vector (Vector). (A) Immunoblotting of the cell lysates with the indicated antibody. (B) Immunoprecipitation of GABARAPL1 with Anti-Flag antibodies covalently bound to agarose followed by immunoblotting with the indicated antibody.

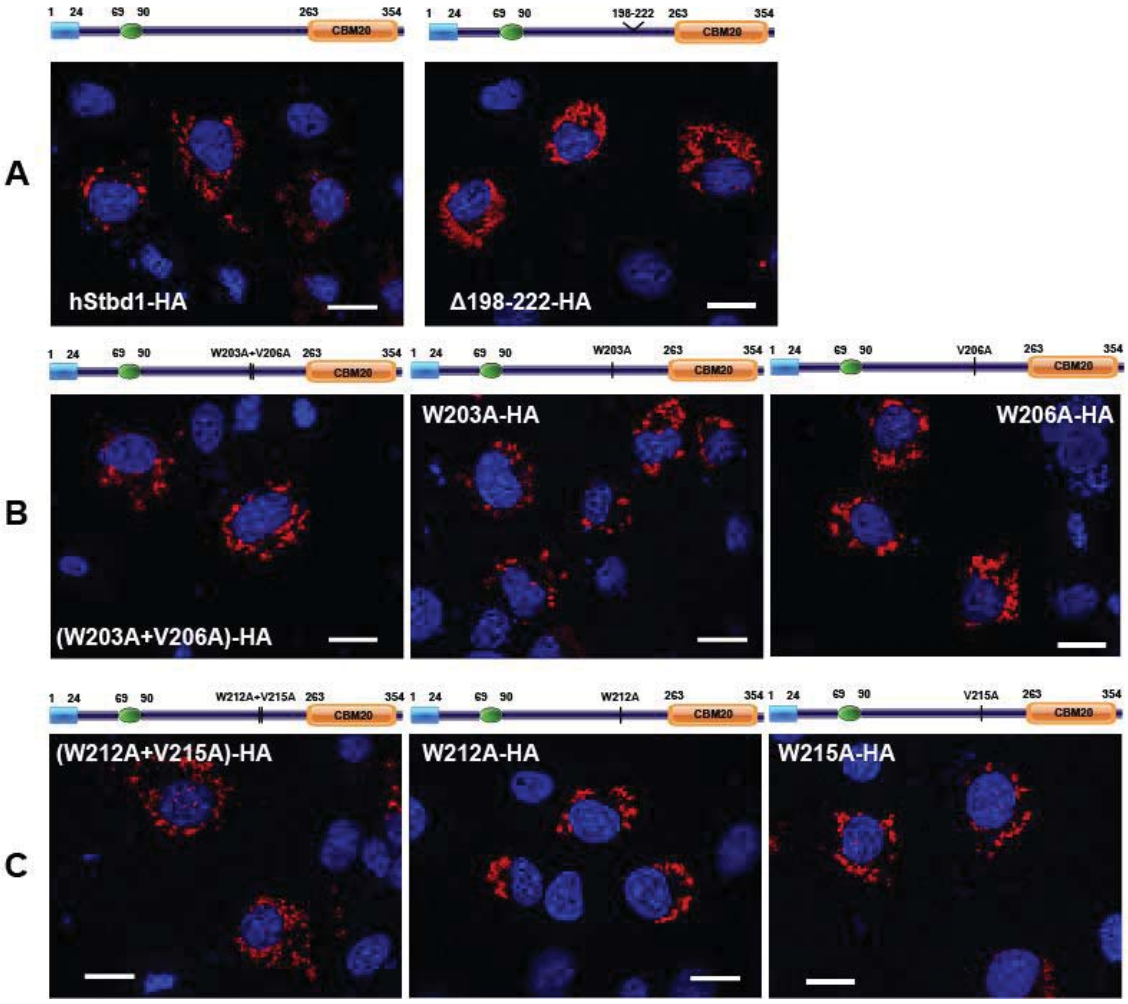


Figure 33. Mutational analysis of Atg8 family interacting motif (AIM) on Stbd1 expressed in COS M9 cells.

Mutated hStbd1 with a C-terminal HA-tag was expressed in COS M9 cells and immunostained with anti-HA antibodies (red). (A) Full length human Stbd1 (left panel) and hStbd1 lacking 198-222 residues (right panel). (B) Mutations at the first potential AIM on Stbd1: double mutation (W203A+V206A-HA) and two single mutations (W203A-HA and V206A-HA). (C) Mutations at the second potential AIM on Stbd1: double mutation (W212A+V215A-HA) and two single mutations (W212A-HA and V215A-HA). Nuclei were stained with Hoechst (blue). The scale bar is 20 μm.

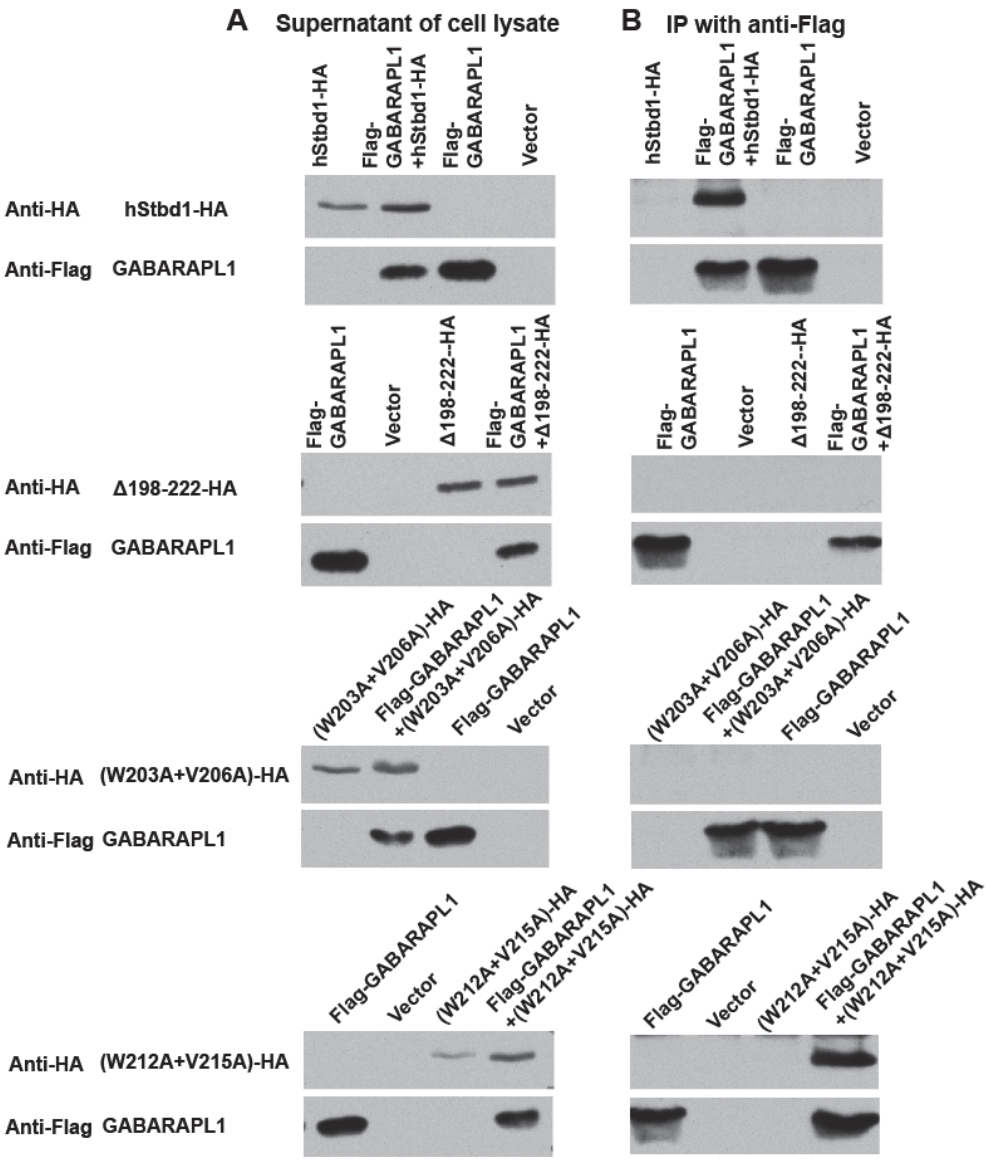


Figure 34. Interaction of GABARAPL1 with Stbd1 and potential Atg8 family interacting motif (AIM) mutants of Stbd1.

hStbd1, full length or mutants lacking 198-222 residues or double point mutations at potential AIM with a C-terminal HA-tag (Δ198-222-HA or W203A+V206A-HA, W212A+V215A-HA) and N-terminal Flag-tagged GABRAPL1 were expressed alone, or in the indicated combination in COS M9 cells. Control cells were transfected with empty pcDNA3 vector (Vector). (A) Immunoblotting of the cell lysates with the indicated antibody (left panels). (B) Immunoprecipitation of GABARAPL1 with Anti-Flag antibodies covalently bound to agarose followed by immunoblotting with the indicated antibody (right panels).

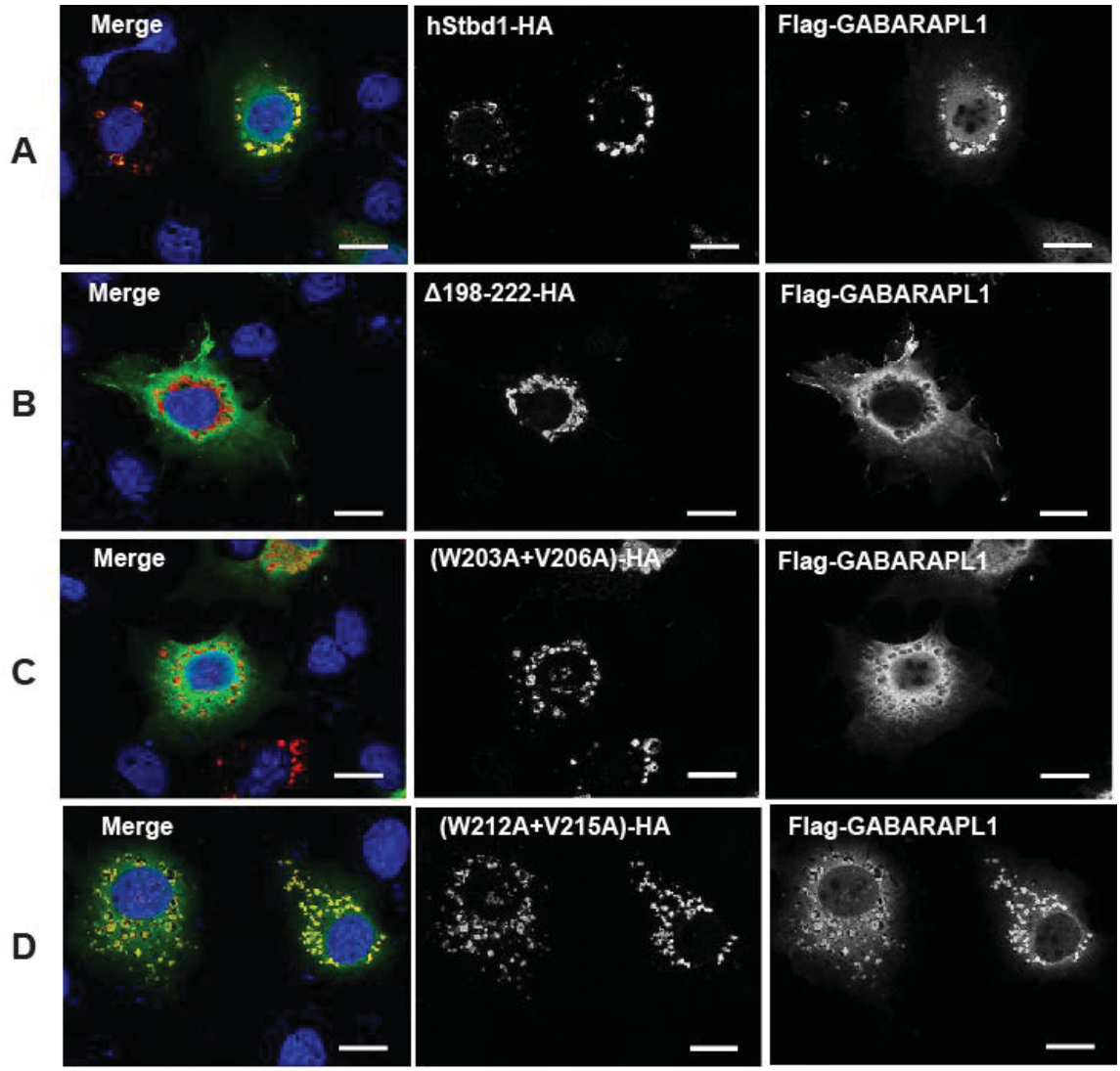


Figure 35. Subcellular localization of GABARAPL1 and Atg8 family interacting motif (AIM) mutants of Stbd1 co-expressed in COS M9 cells. Mutated hStbd1 with a C-terminal HA-tag was co-expressed in COS M9 cells with N-terminal Flag-tagged GABARAPL1 and immunostained with anti-HA antibodies (red) or anti-Flag antibodies (green). (A) Co-localization of hStbd1 and GABARAPL1 (merged in left panel) in cells co-expressing C-terminal HA-tagged full length hStbd1 (middle panel) and N-terminal Flag-tagged GABARAPL1 (right panel). (B) Loss of co-localization (merged in left panel) of Flag-tagged GABARAPL1 (right panel) with potential AIM deletion mutant of hStbd1, Δ198-222-HA (middle panel). (C) Impaired co-localization (merged in left panel) of Flag-tagged GABARAPL1 (right panel) with double mutation in

potential AIM on hStbd1, (W203A+V206A)-HA (middle panel). (D) Unaffected co-localization (merged in left panel) of Flag-tagged GABARAPL1 (right panel) with double mutation in another potential AIM on hStbd1, (W212A+V215A)-HA (middle panel). Nuclei were stained with Hoechst (blue). The scale bar is 20 μm.

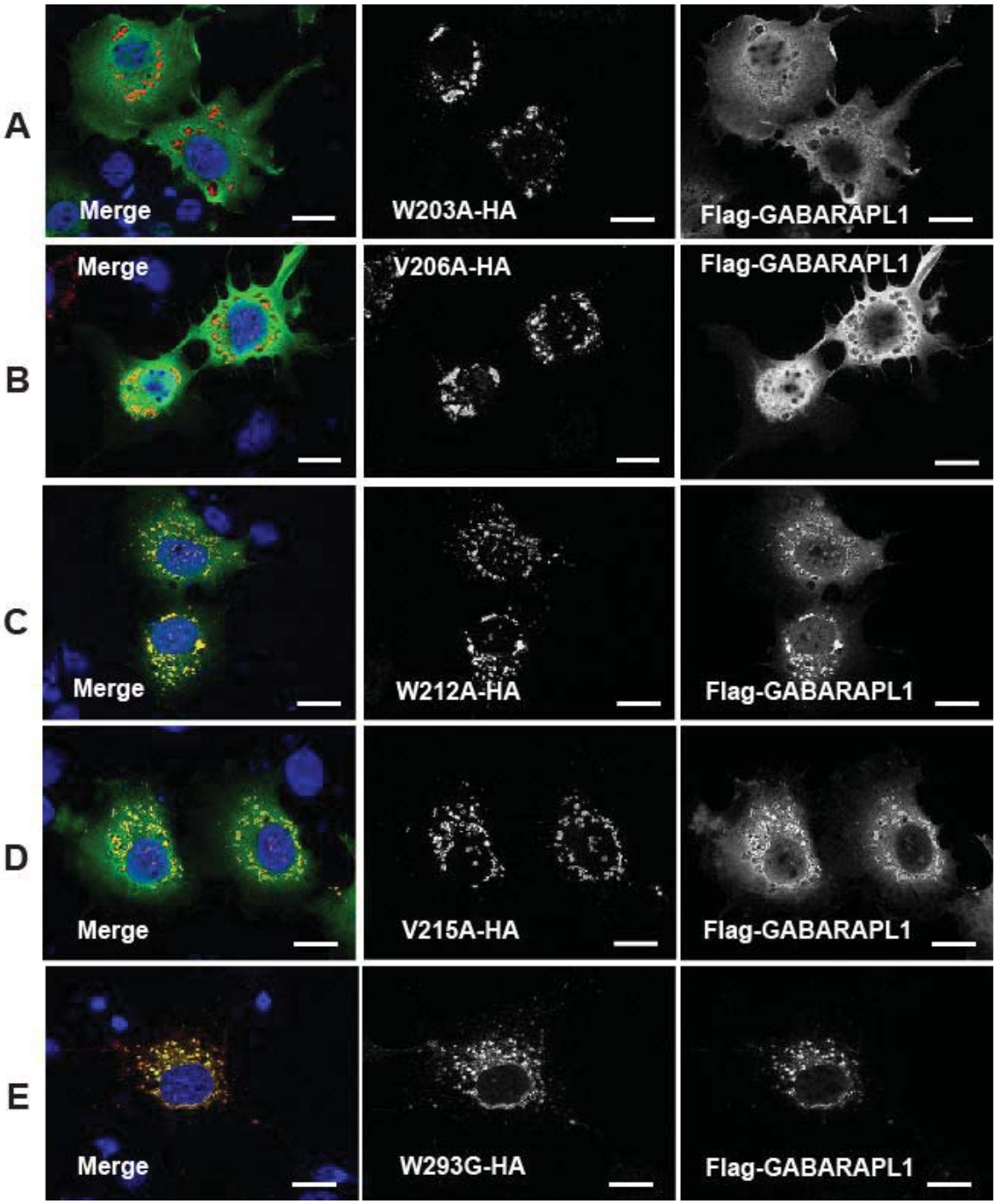


Figure 36. Subcellular localization of GABARAPL1 and single mutants of Stbd1 co-expressed in COS M9 cells.

Mutated hStbd1 with a C-terminal HA-tag was co-expressed in COS M9 cells with N-terminal Flag-tagged GABARAPL1 and immunostained with anti-HA antibodies (red) or anti-Flag antibodies (green). (A) and (B) Impaired co-localization (merged in left panels) of Flag-tagged GABARAPL1 (right panels) with W203A-HA or V206A-HA (middle panels) which contains single mutation within potential AIM on hStbd1. (C) and (D) Unaffected co-localization (merged in left panels) of Flag-tagged GABARAPL1 (right panel) with W212A-HA or V215A-HA (middle panels) which contains single mutation within another potential AIM on hStbd1. (E) Co-localization (merged in left panels) of Flag-tagged GABARAPL1 (right panel) with W293G-HA (middle panels) containing point mutation within CBM20 domain on hStbd1. Nuclei were stained with Hoechst (blue). The scale bar is 20 μm.

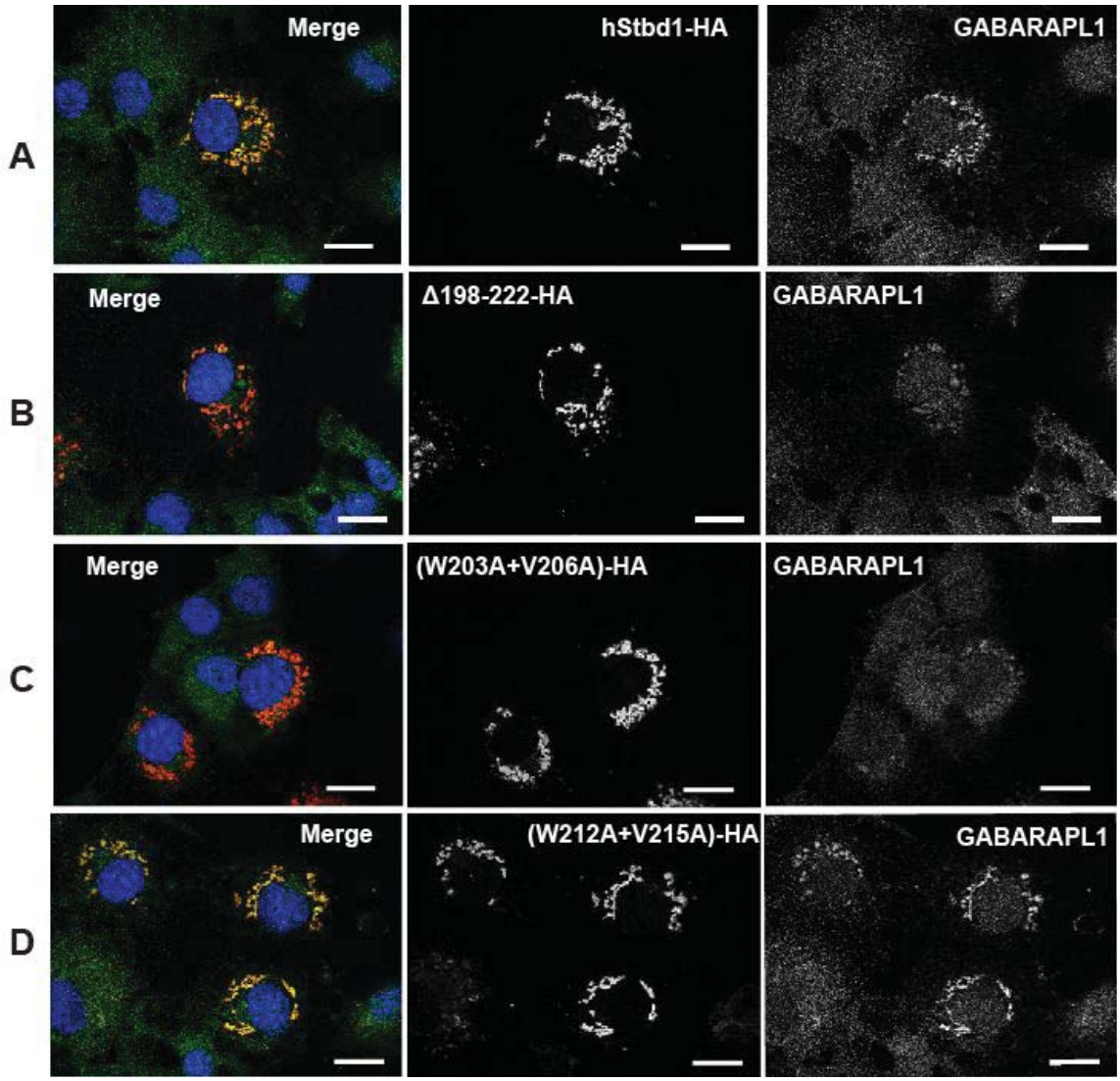


Figure 37. Subcellular localization of endogenous GABARAPL1 and overexpressed Stbd1 with Atg8 family interacting motif (AIM) mutations in COS M9 cells.

Overexpressed full length or mutated hStbd1 with a C-terminal HA-tag and endogenous GABARAPL1 in COS M9 cells was immunostained with anti-HA antibodies (red) or anti-GABARAPL1 antibodies (green). (A) Co-localization of hStbd1 and GABARAPL1 (merged in left panel) in cells expressing C-terminal HA-tagged full length hStbd1 (middle panel) and endogenous GABARAPL1 (right panel). (B) Loss of co-localization (merged in left panel) of endogenous GABARAPL1 (right panel) with potential AIM deletion mutant of hStbd1, Δ198-222-HA (middle panel). (C) Impaired co-localization (merged in left panel) of

endougenous GABARAPL1 (right panel) with double mutation in potential AIM on hStbd1, (W203A+V206A)-HA (middle panel). (D) Unaffected co-localization (merged in left panel) of endougenous GABARAPL1 (right panel) with double mutation in another potential AIM on hStbd1, (W212A+V215A)-HA (middle panel). Nuclei were stained with Hoechst (blue). The scale bar is 20 μm.

DISCUSSION

1. Involvement of Stbd1 in glycogen metabolism

The primary conclusion from this study is that Stbd1 is involved in glycogen metabolism. Data collected in several experiments using independent approaches support this suggestion. We hypothesize that the two most obvious and highly conserved domains of Stbd1, the N-terminal hydrophobic 24 residues and the C-terminal CBM20 domain, serve to anchor glycogen to membranous structures. That the N-terminal hydrophobic segment causes membrane association of Stbd1 is supported by two lines of evidence. First, in muscle extracts, Stbd1 is solubilized from the LSP by non-ionic detergents during fractionation to isolate glycogen. Second, Stbd1 was present in large perinuclear vesicle-like structures and the removal of its N-terminus significantly alters the sub-cellular localization to a diffuse cytosolic distribution. The CBM20 domain at the C-terminus of Stbd1 suggests the ability to bind polysaccharide and is supported by several experiments. First, recombinant Stbd1 co-sedimented with glycogen or the chemically related polysaccharide amylopectin *in vitro*. Second, when non-ionic detergent was present to disrupt membranes when muscle extracts were prepared, there was still a substantial proportion of Stbd1 recovered in the HSP where the glycogen is concentrated, suggesting the co-sedimentation of Stbd1 with glycogen is independent of membrane association. These results suggest that Stbd1 physically interacts with glycogen *in vitro* and in mouse muscle extracts. Third, in the cell models, such as Rat1WT4 cells or COSM9 cells over-expressing Stbd1, where we could visualize accumulated glycogen by PAS staining, Stbd1 was consistently co-localized with glycogen. Further evidence for Stbd1 binding to glycogen in cells was provided by analysis of point mutations of Stbd1 W293, a conserved residue in the CBM20 domain thought to be important for carbohydrate binding. Although over-expressed Stbd1 mutants retained a perinuclear staining pattern for Stbd1, the glycogen was no longer co-localized as judged by PAS staining. The Stbd1-positive

perinuclear structures were not seen at all when Stbd1 lacking the CBM20 domain was expressed. These results are consistent with Stbd1 tethering glycogen to membranes in cells. In other experiments, a genetic link between Stbd1 and glycogen was indicated since Stbd1 protein level decreased when glycogen was genetically reduced in mouse muscle or liver. Other glycogen binding enzymes, such as GS and laforin, have reduced protein levels accompanying genetic depletion of glycogen [86, 228], probably because the binding to glycogen increases the stability of these glycogen associated proteins.

Taking account of these various experimental results, we propose that Stbd1 binds to glycogen by its C-terminal CBM20 domain and interacts with membranes via its hydrophobic N-terminus, thus anchoring glycogen to subcellular membranous structures involving in glycogen metabolism (Figure 31).

Glycogen metabolism can not only be influenced by the regulation of principal metabolic enzymes like GS and GPh, but can also be affected by the glycogen size, subcellular localization and its interaction with numerous proteins of glycogen metabolism [8]. Although glycogen has been studied for decades, there are still many unanswered questions, for instance, where do glycogen particles assemble? The membrane tethering ability of Stbd1 could therefore be involved in determining the location of glycogen synthesis and glycogen particles have been identified in proximity to ER or SR membranes from electronic microscopic data [17, 18] as mentioned in the introduction. Furthermore, in skeletal muscle, the SR localization of glycogen has been suggested as allowing ATP production at sites close to where it is needed for muscular activity [2]. During exercise especially high intensity exercise, continued muscle contractions depend on ATP generated from glycolytic pathway where a large proportion, up to 80% of utilized glucose derives from glycogenolysis. Hence, Stbd1 could assist in localization of glycogen in metabolically useful sites. In fact, in the original paper describing Stbd1, the protein was reported to localize to T-tubules and the sarcoplasmic reticulum of skeletal muscle [161]. Another process that

potentially involves a membrane localization of glycogen is its transport to lysosomes (Figure 2), a process which is clearly demonstrated by the phenotype of GSD-II (Pompe disease) but still not well understood. A notable feature of skeletal muscle pathology in Pompe patients is increased autophagy so that the disease is also categorized as an autophagic vacuolar myopathy [229]. In mouse models of the disease, glycogen is also accumulated in autophagosome-like vesicles or late endosomes as well as lysosomes, which indicates the involvement of intracellular trafficking of glycogen within vesicles [229]. A reasonable explanation for the phenomena in both human patients and mouse models is that defective lysosomal disposal of the glycogen leads to the backing up of cargo vesicles delivering glycogen. Stbd1 could be a critical player in this vesicular trafficking process by tethering polysaccharide molecules to membranes.

2. Stbd1 as a selective autophagic adaptor for glycogen disposal

The participation of Stbd1 in vesicular trafficking of glycogen is also supported by its interaction with GABARAPL1 and to a lesser extent GABARAP which are present in the perinuclear Stbd1-positive and glycogen-enriched structures. Furthermore, Stbd1 bears the specific and conserved sequence, AIM/LIR, which mediates interaction of cargo receptor and Atg8 family proteins in selective autophagy process. Deletion or point mutations of the AIM/LIR in Stbd1 has no effect on Stbd1 subcellular localization, but results in the loss of protein-protein interaction with GABARAPL1, and hence GABARAPL1 is no longer coincident with the perinuclear Stbd1-positive structures. As noted in the Introduction, autophagy can be a selective rather than a totally random process. Depending on target cargo, there are various selective autophagy pathways such as aggrephagy, pexophagy, mitophagy, reticulophagy, ribophagy, lipophagy and xenophagy. To date, only three mammalian proteins have been proposed as autophagic cargo receptors. SQSTM1/p62 and NBR1 mediate trafficking of ubiquitin-modified substrates by binding to LC3. BNIP3L/Nix participates in

mitophagy through two pathways. One is ubiquitin-independent pathway in which BNIP3L/Nix interacting with LC3/GABARAP proteins especially GABARAPL1. In the other pathway, BNIP3L/Nix involves in translocation of E3 ligase Parkin to the membrane of damaged mitochondria and the Parkin-mediated ubiquitination recruits SQSTM1/p62 for subsequent transport [194].

Glycogen autophagy has been discussed in a limited number of reports, where it was observed in skeletal muscle [230], liver and heart [231] of newborn animals that require extensive energy in postnatal metabolism during the period of after birth until the start of suckling [68]. It has been proposed that in newborns, the glycogen autophagy works in addition to the phosphorolytic degradation of glycogen to oppose hypoglycemia. Meanwhile, the process has been suggested to depend on the cAMP-PKA pathway and to be under negative control of the PI-mTOR pathway [67, 68]. However, little is known about glycogen autophagy in adult animals.

We propose that Stbd1 acts as a glycogen receptor interacting with GABARAPL1 and possibly with GABARAP, thereby participating in the autophagic transfer of glycogen to the lysosome (Figure 31). Since the classic autophagosome marker LC3 was not coincident with Stbd1, this vesicular trafficking process is probably distinguished from the canonical autophagy pathway though there may still be involvement of normal macroautophagic machinary. In addition, phospholipid modified GABARAP has been suggested to be a critical element in basal autophagy independent of mTOR inactivation [189]. Therefore, Stbd1-GABARAPL1/GABARAP may play a housekeeping role in cell metabolism, selectively disposing polysaccharides. In our experiments, overexpression of Stbd1 in cells leads to the accumulation of the large perinuclear structures containing glycogen and positive for GABARAPL1 along with Stbd1. These enlarged structures, somewhat similar to what is seen in GSD-II (Pompe disease), may represent accumulated vesicular and glycogen intermediates caused by an overload of the normal trafficking pathway for glycogen

degradation via the lysosome. We also showed in our data that Stbd1 binds better to plant amylopectin or glycogen from *Epm2a* -/- mice, both of which are less branched than normal glycogen, similar to the polyglucosans related to other glycogen storage diseases, such as GSD-IV (Andersen disease or adult polyglucosan disease) and GSD-VII (Tarui disease) [232]. Moreover, in some muscle cells from *Epm2a-/-* and *Epm2b-/-* mice, Stbd1 detected by immunohistochemical staining showed coincidence with Lafora bodies reflected by PAS staining after diastase treatment. Aberrant glycogen obviously need to be removed from cells, hence Stbd1 could selectively target such abnormal polysaccharides for lysosomal clearance via intracellular vesicular delivery. In a proteomic analysis of lysosome-related organelles, Stbd1 was not identified in lysosomes but rather an endosomal fraction instead [233]. This observation would be consistent with lysosomal degradation of Stbd1. In summary, our hypothesis is that glycogen is tethered by Stbd1 to membranes where a vesicle is formed to transport glycogen to lysosomes, with the recruitment of GABARAPL1 or GABARAP, as these vesicles mature prior to lysosomal fusion (Figure 38).

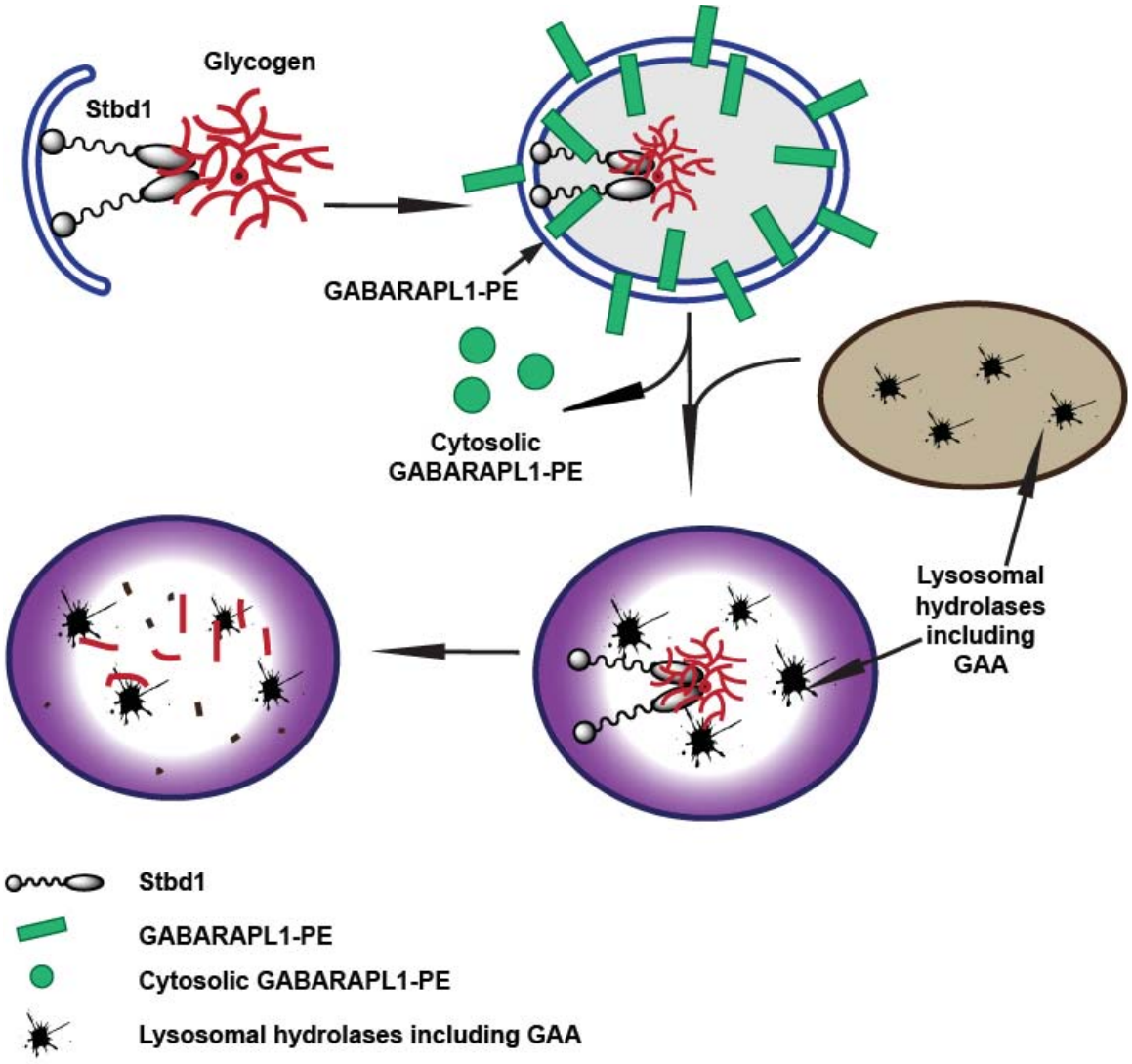


Figure 38. Working model.

Stbd1 acts as a selective autophagic adaptor for glycogen disposal. Stbd1 targets glycogen and tethers it to membranes where a vesicle is formed to transport glycogen to lysosomes for lysosomal degradation, with the recruitment of GABARAPL1 or GABARAP, as these vesicles mature prior to lysosomal fusion.

REFERENCES

1. Roach, P.J. (2002). Glycogen and its metabolism. Curr Mol Med *2*, 101-120.
2. Greenberg, C.C., Jurczak, M.J., Danos, A.M., and Brady, M.J. (2006). Glycogen branches out: new perspectives on the role of glycogen metabolism in the integration of metabolic pathways. Am J Physiol Endocrinol Metab *291*, E1-8.
3. Hartl, P., Olson, E., Dang, T., and Forbes, D.J. (1994). Nuclear assembly with lambda DNA in fractionated Xenopus egg extracts: an unexpected role for glycogen in formation of a higher order chromatin intermediate. J Cell Biol *124*, 235-248.
4. Cheng, C., Mu, J., Farkas, I., Huang, D., Goebl, M.G., and Roach, P.J. (1995). Requirement of the self-glucosylating initiator proteins Glg1p and Glg2p for glycogen accumulation in Saccharomyces cerevisiae. Mol Cell Biol *15*, 6632-6640.
5. Farkas, I., Hardy, T.A., Goebl, M.G., and Roach, P.J. (1991). Two glycogen synthase isoforms in Saccharomyces cerevisiae are coded by distinct genes that are differentially controlled. J Biol Chem *266*, 15602-15607.
6. Ball, S., Guan, H.P., James, M., Myers, A., Keeling, P., Mouille, G., Buleon, A., Colonna, P., and Preiss, J. (1996). From glycogen to amylopectin: a model for the biogenesis of the plant starch granule. Cell *86*, 349-352.
7. Gunja-Smith, Z., Marshall, J.J., Mercier, C., Smith, E.E., and Whelan, W.J. (1970). A revision of the Meyer-Bernfeld model of glycogen and amylopectin. FEBS Lett *12*, 101-104.
8. Shearer, J., and Graham, T.E. (2004). Novel aspects of skeletal muscle glycogen and its regulation during rest and exercise. Exerc Sport Sci Rev *32*, 120-126.
9. Rybicka, K.K. (1996). Glycosomes--the organelles of glycogen metabolism. Tissue Cell *28*, 253-265.
10. Kirkman, B.R., and Whelan, W.J. (1986). Glucosamine is a normal component of liver glycogen. FEBS Lett *194*, 6-11.
11. Fontana, J.D. (1980). The presence of phosphate in glycogen. FEBS Lett *109*, 85-92.
12. Lomako, J., Lomako, W.M., Kirkman, B.R., and Whelan, W.J. (1994). The role of phosphate in muscle glycogen. Biofactors *4*, 167-171.
13. Tarentino, A.L., and Maley, F. (1976). Direct evidence that D-galactosamine incorporation into glycogen occurs via UDP-glucosamine. FEBS Lett *69*, 175-178.
14. Lomako, J., Lomako, W.M., Whelan, W.J., and Marchase, R.B. (1993). Glycogen contains phosphodiester groups that can be introduced by UDPglucose: glycogen glucose 1-phosphotransferase. FEBS Lett *329*, 263-267.

15. Tagliabracci, V.S., Girard, J.M., Segvich, D., Meyer, C., Turnbull, J., Zhao, X., Minassian, B.A., Depaoli-Roach, A.A., and Roach, P.J. (2008). Abnormal metabolism of glycogen phosphate as a cause for Lafora disease. J Biol Chem *283*, 33816-33825.
16. Tagliabracci, V.S., Heiss, C., Karthik, C., Contreras, C.J., Glushka, J., Ishihara, M., Azadi, P., Hurley, T.D., Depaoli-Roach, A.A., and Roach, P.J. (2011). Phosphate Incorporation during Glycogen Synthesis and Lafora Disease. Cell Metab *13*, 274-282.
17. Cardell, R.R., Jr., Michaels, J.E., Hung, J.T., and Cardell, E.L. (1985). SERGE, the subcellular site of initial hepatic glycogen deposition in the rat: a radioautographic and cytochemical study. J Cell Biol *101*, 201-206.
18. Meyer, F., Heilmeyer, L.M., Jr., Haschke, R.H., and Fischer, E.H. (1970). Control of phosphorylase activity in a muscle glycogen particle. I. Isolation and characterization of the protein-glycogen complex. J Biol Chem *245*, 6642-6648.
19. Takeshige, K., Baba, M., Tsuboi, S., Noda, T., and Ohsumi, Y. (1992). Autophagy in yeast demonstrated with proteinase-deficient mutants and conditions for its induction. J Cell Biol *119*, 301-311.
20. Geddes, R., and Stratton, G.C. (1977). The influence of lysosomes on glycogen metabolism. Biochem J *163*, 193-200.
21. Geddes, R. (1986). Glycogen: a metabolic viewpoint. Biosci Rep *6*, 415-428.
22. Olson, A.L., and Pessin, J.E. (1996). Structure, function, and regulation of the mammalian facilitative glucose transporter gene family. Annu Rev Nutr *16*, 235-256.
23. Thorens, B. (1996). Glucose transporters in the regulation of intestinal, renal, and liver glucose fluxes. Am J Physiol *270*, G541-553.
24. Simpson, F., Whitehead, J.P., and James, D.E. (2001). GLUT4--at the cross roads between membrane trafficking and signal transduction. Traffic *2*, 2-11.
25. Cori, C.F.C.G.T. (1939). The activating effect of glycogen on the enzymatic synthesis of glycogen from glucose-1-phosphate. JBC *131*, 397-398.
26. Leloir, L.F., Olavarria, J.M., Goldemberg, S.H., and Carminatti, H. (1959). Biosynthesis of glycogen from uridine diphosphate glucose. Arch Biochem Biophys *81*, 508-520.
27. Krisman, C.R., and Barengo, R. (1975). A precursor of glycogen biosynthesis: alpha-1,4-glucan-protein. Eur J Biochem *52*, 117-123.
28. Butler, N.A., Lee, E.Y., and Whelan, W.J. (1977). A protein-bound glycogen component of rat liver. Carbohydr Res *55*, 73-82.
29. Gibbons, B.J., Roach, P.J., and Hurley, T.D. (2002). Crystal structure of the autocatalytic initiator of glycogen biosynthesis, glycogenin. J Mol Biol *319*, 463-477.
30. Moslemi, A.R., Lindberg, C., Nilsson, J., Tajsharghi, H., Andersson, B., and Oldfors, A. (2010). Glycogenin-1 deficiency and inactivated priming of glycogen synthesis. N Engl J Med *362*, 1203-1210.

31. Mu, J., Skurat, A.V., and Roach, P.J. (1997). Glycogenin-2, a novel self-glucosylating protein involved in liver glycogen biosynthesis. J Biol Chem *272*, 27589-27597.
32. Rodriguez, I.R., and Whelan, W.J. (1985). A novel glycosyl-amino acid linkage: rabbit-muscle glycogen is covalently linked to a protein via tyrosine. Biochem Biophys Res Commun *132*, 829-836.
33. Smythe, C., Caudwell, F.B., Ferguson, M., and Cohen, P. (1988). Isolation and structural analysis of a peptide containing the novel tyrosyl-glucose linkage in glycogenin. EMBO J *7*, 2681-2686.
34. Lomako, J., Lomako, W.M., and Whelan, W.J. (1988). A self-glucosylating protein is the primer for rabbit muscle glycogen biosynthesis. FASEB J *2*, 3097-3103.
35. Hurley, T.D., Walls, C., Bennett, J.R., Roach, P.J., and Wang, M. (2006). Direct detection of glycogenin reaction products during glycogen initiation. Biochem Biophys Res Commun *348*, 374-378.
36. Pederson, B.A., Cheng, C., Wilson, W.A., and Roach, P.J. (2000). Regulation of glycogen synthase. Identification of residues involved in regulation by the allosteric ligand glucose-6-P and by phosphorylation. J Biol Chem *275*, 27753-27761.
37. Friedman, D.L., and Larner, J. (1963). Studies on Udpg-Alpha-Glucan Transglucosylase. Iii. Interconversion of Two Forms of Muscle Udpg-Alpha-Glucan Transglucosylase by a Phosphorylation-Dephosphorylation Reaction Sequence. Biochemistry *2*, 669-675.
38. Roach, R.J., and Larner, J. (1977). Covalent phosphorylation in the regulation glycogen synthase activity. Mol Cell Biochem *15*, 179-200.
39. Picton, C., Woodgett, J., Hemmings, B., and Cohen, P. (1982). Multisite phosphorylation of glycogen synthase from rabbit skeletal muscle. Phosphorylation of site 5 by glycogen synthase kinase-5 (casein kinase-II) is a prerequisite for phosphorylation of sites 3 by glycogen synthase kinase-3. FEBS Lett *150*, 191-196.
40. DePaoli-Roach, A.A., Ahmad, Z., Camici, M., Lawrence, J.C., Jr., and Roach, P.J. (1983). Multiple phosphorylation of rabbit skeletal muscle glycogen synthase. Evidence for interactions among phosphorylation sites and the resolution of electrophoretically distinct forms of the subunit. J Biol Chem *258*, 10702-10709.
41. Roach, P.J., DePaoli-Roach, A.A., and Larner, J. (1978). Ca2+-stimulated phosphorylation of muscle glycogen synthase by phosphorylase b kinase. J Cyclic Nucleotide Res *4*, 245-257.
42. Huang, T.S., and Krebs, E.G. (1977). Amino acid sequence of a phosphorylation site in skeletal muscle glycogen synthetase. Biochem Biophys Res Commun *75*, 643-650.
43. Flotow, H., and Roach, P.J. (1989). Synergistic phosphorylation of rabbit muscle glycogen synthase by cyclic AMP-dependent protein kinase and casein kinase I. Implications for hormonal regulation of glycogen synthase. J Biol Chem *264*, 9126-9128.

44. Flotow, H., Graves, P.R., Wang, A.Q., Fiol, C.J., Roeske, R.W., and Roach, P.J. (1990). Phosphate groups as substrate determinants for casein kinase I action. J Biol Chem *265*, 14264-14269.
45. Fiol, C.J., Mahrenholz, A.M., Wang, Y., Roeske, R.W., and Roach, P.J. (1987). Formation of protein kinase recognition sites by covalent modification of the substrate. Molecular mechanism for the synergistic action of casein kinase II and glycogen synthase kinase 3. J Biol Chem *262*, 14042-14048.
46. Carling, D., and Hardie, D.G. (1989). The substrate and sequence specificity of the AMP-activated protein kinase. Phosphorylation of glycogen synthase and phosphorylase kinase. Biochim Biophys Acta *1012*, 81-86.
47. Wilson, W.A., Skurat, A.V., Probst, B., de Paoli-Roach, A., Roach, P.J., and Rutter, J. (2005). Control of mammalian glycogen synthase by PAS kinase. Proc Natl Acad Sci U S A *102*, 16596-16601.
48. Skurat, A.V., and Dietrich, A.D. (2004). Phosphorylation of Ser640 in muscle glycogen synthase by DYRK family protein kinases. J Biol Chem *279*, 2490-2498.
49. Kuma, Y., Campbell, D.G., and Cuenda, A. (2004). Identification of glycogen synthase as a new substrate for stress-activated protein kinase 2b/p38beta. Biochem J *379*, 133-139.
50. Bollen, M., Peti, W., Ragusa, M.J., and Beullens, M. (2010). The extended PP1 toolkit: designed to create specificity. Trends Biochem Sci *35*, 450-458.
51. Ceulemans, H., Stalmans, W., and Bollen, M. (2002). Regulator-driven functional diversification of protein phosphatase-1 in eukaryotic evolution. Bioessays *24*, 371-381.
52. Ceulemans, H., and Bollen, M. (2004). Functional diversity of protein phosphatase-1, a cellular economizer and reset button. Physiol Rev *84*, 1-39.
53. Stralfors, P., Hiraga, A., and Cohen, P. (1985). The protein phosphatases involved in cellular regulation. Purification and characterisation of the glycogen-bound form of protein phosphatase-1 from rabbit skeletal muscle. Eur J Biochem *149*, 295-303.
54. Aschenbach, W.G., Suzuki, Y., Breeden, K., Prats, C., Hirshman, M.F., Dufresne, S.D., Sakamoto, K., Vilardo, P.G., Steele, M., Kim, J.H., et al. (2001). The muscle-specific protein phosphatase PP1G/R(GL)(G(M))is essential for activation of glycogen synthase by exercise. J Biol Chem *276*, 39959-39967.
55. Savage, D.B., Zhai, L., Ravikumar, B., Choi, C.S., Snaar, J.E., McGuire, A.C., Wou, S.E., Medina-Gomez, G., Kim, S., Bock, C.B., et al. (2008). A prevalent variant in PPP1R3A impairs glycogen synthesis and reduces muscle glycogen content in humans and mice. PLoS Med *5*, e27.

56. Munro, S., Cuthbertson, D.J., Cunningham, J., Sales, M., and Cohen, P.T. (2002). Human skeletal muscle expresses a glycogen-targeting subunit of PP1 that is identical to the insulin-sensitive glycogen-targeting subunit G(L) of liver. Diabetes *51*, 591-598.
57. Stalmans, W., Cadefau, J., Wera, S., and Bollen, M. (1997). New insight into the regulation of liver glycogen metabolism by glucose. Biochem Soc Trans *25*, 19-25.
58. Alemany, S., and Cohen, P. (1986). Phosphorylase a is an allosteric inhibitor of the glycogen and microsomal forms of rat hepatic protein phosphatase-1. FEBS Lett *198*, 194-202.
59. Doherty, M.J., Young, P.R., and Cohen, P.T. (1996). Amino acid sequence of a novel protein phosphatase 1 binding protein (R5) which is related to the liver- and muscle-specific glycogen binding subunits of protein phosphatase 1. FEBS Lett *399*, 339-343.
60. Printen, J.A., Brady, M.J., and Saltiel, A.R. (1997). PTG, a protein phosphatase 1-binding protein with a role in glycogen metabolism. Science *275*, 1475-1478.
61. Brady, M.J., Printen, J.A., Mastick, C.C., and Saltiel, A.R. (1997). Role of protein targeting to glycogen (PTG) in the regulation of protein phosphatase-1 activity. J Biol Chem *272*, 20198-20204.
62. Armstrong, C.G., Browne, G.J., Cohen, P., and Cohen, P.T. (1997). PPP1R6, a novel member of the family of glycogen-targetting subunits of protein phosphatase 1. FEBS Lett *418*, 210-214.
63. Munro, S., Ceulemans, H., Bollen, M., Diplexcito, J., and Cohen, P.T. (2005). A novel glycogen-targeting subunit of protein phosphatase 1 that is regulated by insulin and shows differential tissue distribution in humans and rodents. FEBS J *272*, 1478-1489.
64. Gibson, W.B., Illingsworth, B., and Brown, D.H. (1971). Studies of glycogen branching enzyme. Preparation and properties of -1,4-glucan- -1,4-glucan 6-glycosyltransferase and its action on the characteristic polysaccharide of the liver of children with Type IV glycogen storage disease. Biochemistry *10*, 4253-4262.
65. Caudwell, F.B., and Cohen, P. (1980). Purification and subunit structure of glycogen-branching enzyme from rabbit skeletal muscle. Eur J Biochem *109*, 391-394.
66. Raben, N., Danon, M., Lu, N., Lee, E., Shliselfeld, L., Skurat, A.V., Roach, P.J., Lawrence, J.C., Jr., Musumeci, O., Shanske, S., et al. (2001). Surprises of genetic engineering: a possible model of polyglucosan body disease. Neurology *56*, 1739-1745.
67. Kotoulas, O.B., Kalamidas, S.A., and Kondomerkos, D.J. (2004). Glycogen autophagy. Microsc Res Tech *64*, 10-20.
68. Kotoulas, O.B., Kalamidas, S.A., and Kondomerkos, D.J. (2006). Glycogen autophagy in glucose homeostasis. Pathol Res Pract *202*, 631-638.

69. Newgard, C.B., Hwang, P.K., and Fletterick, R.J. (1989). The family of glycogen phosphorylases: structure and function. Crit Rev Biochem Mol Biol *24*, 69-99.
70. Johnson, L.N. (1992). Glycogen phosphorylase: control by phosphorylation and allosteric effectors. FASEB J *6*, 2274-2282.
71. Titani, K., Koide, A., Ericsson, L.H., Kumar, S., Hermann, J., Wade, R.D., Walsh, K.A., Neurath, H., and Fischer, E.H. (1978). Sequence of the carboxyl-terminal 492 residues of rabbit muscle glycogen phosphorylase including the pyridoxal 5'-phosphate binding site. Biochemistry *17*, 5680-5693.
72. Cohen, P.T. (2002). Protein phosphatase 1--targeted in many directions. J Cell Sci *115*, 241-256.
73. Gilboe, D.P., Larson, K.L., and Nuttall, F.Q. (1972). Radioactive method for the assay of glycogen phosphorylases. Anal Biochem *47*, 20-27.
74. Taylor, C., Cox, A.J., Kernohan, J.C., and Cohen, P. (1975). Debranching enzyme from rabbit skeletal muscle. Purification, properties and physiological role. Eur J Biochem *51*, 105-115.
75. Bates, E.J., Heaton, G.M., Taylor, C., Kernohan, J.C., and Cohen, P. (1975). Debranching enzyme from rabbit skeletal muscle; evidence for the location of two active centres on a single polypeptide chain. FEBS Lett *58*, 181-185.
76. Henrissat, B., and Davies, G.J. (2000). Glycoside hydrolases and glycosyltransferases. Families, modules, and implications for genomics. Plant Physiol *124*, 1515-1519.
77. Wisselaar, H.A., Kroos, M.A., Hermans, M.M., van Beeumen, J., and Reuser, A.J. (1993). Structural and functional changes of lysosomal acid alpha-glucosidase during intracellular transport and maturation. J Biol Chem *268*, 2223-2231.
78. Hermans, M.M., Wisselaar, H.A., Kroos, M.A., Oostra, B.A., and Reuser, A.J. (1993). Human lysosomal alpha-glucosidase: functional characterization of the glycosylation sites. Biochem J *289 (Pt 3)*, 681-686.
79. Hauser, H., and Semenza, G. (1983). Sucrase-isomaltase: a stalked intrinsic protein of the brush border membrane. CRC Crit Rev Biochem *14*, 319-345.
80. Hermans, M.M., Kroos, M.A., van Beeumen, J., Oostra, B.A., and Reuser, A.J. (1991). Human lysosomal alpha-glucosidase. Characterization of the catalytic site. J Biol Chem *266*, 13507-13512.
81. Yan, B., Heus, J., Lu, N., Nichols, R.C., Raben, N., and Plotz, P.H. (2001). Transcriptional regulation of the human acid alpha-glucosidase gene. Identification of a repressor element and its transcription factors Hes-1 and YY1. J Biol Chem *276*, 1789-1793.
82. Lawrence, J.C., Jr., Hiken, J.F., DePaoli-Roach, A.A., and Roach, P.J. (1983). Hormonal control of glycogen synthase in rat hemidiaphragms. Effects of insulin and epinephrine on the distribution of phosphate between two cyanogen bromide fragments. J Biol Chem *258*, 10710-10719.

83. Parker, P.J., Caudwell, F.B., and Cohen, P. (1983). Glycogen synthase from rabbit skeletal muscle; effect of insulin on the state of phosphorylation of the seven phosphoserine residues in vivo. Eur J Biochem *130*, 227-234.
84. Dent, P., Lavoinne, A., Nakielny, S., Caudwell, F.B., Watt, P., and Cohen, P. (1990). The molecular mechanism by which insulin stimulates glycogen synthesis in mammalian skeletal muscle. Nature *348*, 302-308.
85. Brady, M.J., and Saltiel, A.R. (2001). The role of protein phosphatase-1 in insulin action. Recent Prog Horm Res *56*, 157-173.
86. Suzuki, Y., Lanner, C., Kim, J.H., Vilardo, P.G., Zhang, H., Yang, J., Cooper, L.D., Steele, M., Kennedy, A., Bock, C.B., et al. (2001). Insulin control of glycogen metabolism in knockout mice lacking the muscle-specific protein phosphatase PP1G/RGL. Mol Cell Biol *21*, 2683-2694.
87. MacKintosh, C., Campbell, D.G., Hiraga, A., and Cohen, P. (1988). Phosphorylation of the glycogen-binding subunit of protein phosphatase-1G in response to adrenalin. FEBS Lett *234*, 189-194.
88. Foulkes, J.G., and Cohen, P. (1979). The hormonal control of glycogen metabolism. Phosphorylation of protein phosphatase inhibitor-1 in vivo in response to adrenaline. Eur J Biochem *97*, 251-256.
89. Cohen, P. (1989). The structure and regulation of protein phosphatases. Annu Rev Biochem *58*, 453-508.
90. Moorhead, G., MacKintosh, C., Morrice, N., and Cohen, P. (1995). Purification of the hepatic glycogen-associated form of protein phosphatase-1 by microcystin-Sepharose affinity chromatography. FEBS Lett *362*, 101-105.
91. Lewis, G.M., Spencer-Peet, J., and Stewart, K.M. (1963). Infantile Hypoglycaemia due to Inherited Deficiency of Glycogen Synthetase in Liver. Arch Dis Child *38*, 40-48.
92. Orho, M., Bosshard, N.U., Buist, N.R., Gitzelmann, R., Aynsley-Green, A., Blumel, P., Gannon, M.C., Nuttall, F.Q., and Groop, L.C. (1998). Mutations in the liver glycogen synthase gene in children with hypoglycemia due to glycogen storage disease type 0. J Clin Invest *102*, 507-515.
93. Irimia, J.M., Meyer, C.M., Peper, C.L., Zhai, L., Bock, C.B., Previs, S.F., McGuinness, O.P., DePaoli-Roach, A., and Roach, P.J. (2010). Impaired glucose tolerance and predisposition to the fasted state in liver glycogen synthase knock-out mice. J Biol Chem *285*, 12851-12861.
94. Cameron, J.M., Levandovskiy, V., MacKay, N., Utgikar, R., Ackerley, C., Chiasson, D., Halliday, W., Raiman, J., and Robinson, B.H. (2009). Identification of a novel mutation in GYS1 (muscle-specific glycogen synthase) resulting in sudden cardiac death, that is diagnosable from skin fibroblasts. Mol Genet Metab *98*, 378-382.
95. Pederson, B.A., Chen, H., Schroeder, J.M., Shou, W., DePaoli-Roach, A.A., and Roach, P.J. (2004). Abnormal cardiac development in the absence of heart glycogen. Mol Cell Biol *24*, 7179-7187.

96. Koeberl, D.D., Kishnani, P.S., Bali, D., and Chen, Y.T. (2009). Emerging therapies for glycogen storage disease type I. Trends Endocrinol Metab *20*, 252-258.
97. Chou, J.Y., Matern, D., Mansfield, B.C., and Chen, Y.T. (2002). Type I glycogen storage diseases: disorders of the glucose-6-phosphatase complex. Curr Mol Med *2*, 121-143.
98. Huijing, F. (1975). Glycogen metabolism and glycogen-storage diseases. Physiol Rev *55*, 609-658.
99. Ozen, H. (2007). Glycogen storage diseases: new perspectives. World J Gastroenterol *13*, 2541-2553.
100. Melis, D., Havelaar, A.C., Verbeek, E., Smit, G.P., Benedetti, A., Mancini, G.M., and Verheijen, F. (2004). NPT4, a new microsomal phosphate transporter: mutation analysis in glycogen storage disease type Ic. J Inherit Metab Dis *27*, 725-733.
101. Raben, N., Plotz, P., and Byrne, B.J. (2002). Acid alpha-glucosidase deficiency (glycogenosis type II, Pompe disease). Curr Mol Med *2*, 145-166.
102. Slonim, A.E., Bulone, L., Ritz, S., Goldberg, T., Chen, A., and Martiniuk, F. (2000). Identification of two subtypes of infantile acid maltase deficiency. J Pediatr *137*, 283-285.
103. Martin, J.J., De Barsy, T., De, S., Leroy, J.G., and Palladini, G. (1976). Acid maltase deficiency (type II glycogenosis). Morphological and biochemical study of a childhood phenotype. J Neurol Sci *30*, 155-166.
104. Ausems, M.G., Lochman, P., van Diggelen, O.P., Ploos van Amstel, H.K., Reuser, A.J., and Wokke, J.H. (1999). A diagnostic protocol for adult-onset glycogen storage disease type II. Neurology *52*, 851-853.
105. Raben, N., Nagaraju, K., Lee, E., Kessler, P., Byrne, B., Lee, L., LaMarca, M., King, C., Ward, J., Sauer, B., et al. (1998). Targeted disruption of the acid alpha-glucosidase gene in mice causes an illness with critical features of both infantile and adult human glycogen storage disease type II. J Biol Chem *273*, 19086-19092.
106. Douillard-Guilloux, G., Raben, N., Takikita, S., Ferry, A., Vignaud, A., Guillet-Deniau, I., Favier, M., Thurberg, B.L., Roach, P.J., Caillaud, C., et al. (2010). Restoration of muscle functionality by genetic suppression of glycogen synthesis in a murine model of Pompe disease. Hum Mol Genet *19*, 684-696.
107. Shen, J.J., and Chen, Y.T. (2002). Molecular characterization of glycogen storage disease type III. Curr Mol Med *2*, 167-175.
108. Talente, G.M., Coleman, R.A., Alter, C., Baker, L., Brown, B.I., Cannon, R.A., Chen, Y.T., Crigler, J.F., Jr., Ferreira, P., Haworth, J.C., et al. (1994). Glycogen storage disease in adults. Ann Intern Med *120*, 218-226.
109. Shen, J., Bao, Y., Liu, H.M., Lee, P., Leonard, J.V., and Chen, Y.T. (1996). Mutations in exon 3 of the glycogen debranching enzyme gene are associated with glycogen storage disease type III that is differentially expressed in liver and muscle. J Clin Invest *98*, 352-357.

110. Van Hoof, F., and Hers, H.G. (1967). The subgroups of type 3 glycogenosis. Eur J Biochem *2*, 265-270.
111. Ding, J.H., de Barsy, T., Brown, B.I., Coleman, R.A., and Chen, Y.T. (1990). Immunoblot analyses of glycogen debranching enzyme in different subtypes of glycogen storage disease type III. J Pediatr *116*, 95-100.
112. Wolfsdorf, J.I., and Weinstein, D.A. (2003). Glycogen storage diseases. Rev Endocr Metab Disord *4*, 95-102.
113. Andersen, D.H. (1956). Familial cirrhosis of the liver with storage of abnormal glycogen. Lab Invest *5*, 11-20.
114. Schroder, J.M., May, R., Shin, Y.S., Sigmund, M., and Nase-Huppmeier, S. (1993). Juvenile hereditary polyglucosan body disease with complete branching enzyme deficiency (type IV glycogenosis). Acta Neuropathol *85*, 419-430.
115. Raju, G.P., Li, H.C., Bali, D.S., Chen, Y.T., Urion, D.K., Lidov, H.G., and Kang, P.B. (2008). A case of congenital glycogen storage disease type IV with a novel GBE1 mutation. J Child Neurol *23*, 349-352.
116. Dimaur, S., Andreu, A.L., Bruno, C., and Hadjigeorgiou, G.M. (2002). Myophosphorylase deficiency (glycogenosis type V; McArdle disease). Curr Mol Med *2*, 189-196.
117. Nakajima, H., Raben, N., Hamaguchi, T., and Yamasaki, T. (2002). Phosphofructokinase deficiency; past, present and future. Curr Mol Med *2*, 197-212.
118. Davidson, J.J., Ozcelik, T., Hamacher, C., Willems, P.J., Francke, U., and Kilimann, M.W. (1992). cDNA cloning of a liver isoform of the phosphorylase kinase alpha subunit and mapping of the gene to Xp22.2-p22.1, the region of human X-linked liver glycogenosis. Proc Natl Acad Sci U S A *89*, 2096-2100.
119. Hendrickx, J., and Willems, P.J. (1996). Genetic deficiencies of the glycogen phosphorylase system. Hum Genet *97*, 551-556.
120. Burwinkel, B., Maichele, A.J., Aagenaes, O., Bakker, H.D., Lerner, A., Shin, Y.S., Strachan, J.A., and Kilimann, M.W. (1997). Autosomal glycogenosis of liver and muscle due to phosphorylase kinase deficiency is caused by mutations in the phosphorylase kinase beta subunit (PHKB). Hum Mol Genet *6*, 1109-1115.
121. Burwinkel, B., Shiomi, S., Al Zaben, A., and Kilimann, M.W. (1998). Liver glycogenosis due to phosphorylase kinase deficiency: PHKG2 gene structure and mutations associated with cirrhosis. Hum Mol Genet *7*, 149-154.
122. Ganesh, S., Puri, R., Singh, S., Mittal, S., and Dubey, D. (2006). Recent advances in the molecular basis of Lafora's progressive myoclonus epilepsy. J Hum Genet *51*, 1-8.
123. Chan, E.M., Andrade, D.M., Franceschetti, S., and Minassian, B. (2005). Progressive myoclonus epilepsies: EPM1, EPM2A, EPM2B. Adv Neurol *95*, 47-57.

124. Chan, E.M., Omer, S., Ahmed, M., Bridges, L.R., Bennett, C., Scherer, S.W., and Minassian, B.A. (2004). Progressive myoclonus epilepsy with polyglucosans (Lafora disease): evidence for a third locus. Neurology *63*, 565-567.
125. Singh, S., and Ganesh, S. (2009). Lafora progressive myoclonus epilepsy: a meta-analysis of reported mutations in the first decade following the discovery of the EPM2A and NHLRC1 genes. Hum Mutat *30*, 715-723.
126. Monaghan, T.S., and Delanty, N. (2010). Lafora disease: epidemiology, pathophysiology and management. CNS Drugs *24*, 549-561.
127. Shoseyov, O., Shani, Z., and Levy, I. (2006). Carbohydrate binding modules: biochemical properties and novel applications. Microbiol Mol Biol Rev *70*, 283-295.
128. Hashimoto, H. (2006). Recent structural studies of carbohydrate-binding modules. Cell Mol Life Sci *63*, 2954-2967.
129. Gilkes, N.R., Warren, R.A., Miller, R.C., Jr., and Kilburn, D.G. (1988). Precise excision of the cellulose binding domains from two Cellulomonas fimi cellulases by a homologous protease and the effect on catalysis. J Biol Chem *263*, 10401-10407.
130. Tomme, P., Van Tilbeurgh, H., Pettersson, G., Van Damme, J., Vandekerckhove, J., Knowles, J., Teeri, T., and Claeyssens, M. (1988). Studies of the cellulolytic system of Trichoderma reesei QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis. Eur J Biochem *170*, 575-581.
131. Boraston, A.B., Bolam, D.N., Gilbert, H.J., and Davies, G.J. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem J *382*, 769-781.
132. Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res *37*, D233-238.
133. Guillen, D., Sanchez, S., and Rodriguez-Sanoja, R. (2010). Carbohydrate-binding domains: multiplicity of biological roles. Appl Microbiol Biotechnol *85*, 1241-1249.
134. Richardson, J.S. (1981). The anatomy and taxonomy of protein structure. Adv Protein Chem *34*, 167-339.
135. Murzin, A.G., Lesk, A.M., and Chothia, C. (1992). beta-Trefoil fold. Patterns of structure and sequence in the Kunitz inhibitors interleukins-1 beta and 1 alpha and fibroblast growth factors. J Mol Biol *223*, 531-543.
136. Murzin, A.G. (1993). OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. EMBO J *12*, 861-867.

137. Quiocho, F.A. (1986). Carbohydrate-binding proteins: tertiary structures and protein-sugar interactions. Annu Rev Biochem *55*, 287-315.
138. Barral, P., Suarez, C., Batanero, E., Alfonso, C., Alche Jde, D., Rodriguez-Garcia, M.I., Villalba, M., Rivas, G., and Rodriguez, R. (2005). An olive pollen protein with allergenic activity, Ole e 10, defines a novel family of carbohydrate-binding modules and is potentially implicated in pollen germination. Biochem J *390*, 77-84.
139. Vaaje-Kolstad, G., Horn, S.J., van Aalten, D.M., Synstad, B., and Eijsink, V.G. (2005). The non-catalytic chitin-binding protein CBP21 from Serratia marcescens is essential for chitin degradation. J Biol Chem *280*, 28492-28497.
140. Moser, F., Irwin, D., Chen, S., and Wilson, D.B. (2008). Regulation and characterization of Thermobifida fusca carbohydrate-binding module proteins E7 and E8. Biotechnol Bioeng *100*, 1066-1077.
141. Machovic, M., and Janecek, S. (2006). Starch-binding domains in the post-genome era. Cell Mol Life Sci *63*, 2710-2724.
142. Machovic, M., Svensson, B., MacGregor, E.A., and Janecek, S. (2005). A new clan of CBM families based on bioinformatics of starch-binding domains from families CBM20 and CBM21. FEBS J *272*, 5497-5513.
143. Christiansen, C., Hachem, M.A., Glaring, M.A., Vikso-Nielsen, A., Sigurskjold, B.W., Svensson, B., and Blennow, A. (2009). A CBM20 low-affinity starch-binding domain from glucan, water dikinase. FEBS Lett *583*, 1159-1163.
144. Machovic, M., and Janecek, S. (2006). The evolution of putative starch-binding domains. FEBS Lett *580*, 6349-6356.
145. Lawson, C.L., van Montfort, R., Strokopytov, B., Rozeboom, H.J., Kalk, K.H., de Vries, G.E., Penninga, D., Dijkhuizen, L., and Dijkstra, B.W. (1994). Nucleotide sequence and X-ray structure of cyclodextrin glycosyltransferase from Bacillus circulans strain 251 in a maltose-dependent crystal form. J Mol Biol *236*, 590-600.
146. Sorimachi, K., Jacks, A.J., Le Gal-Coeffet, M.F., Williamson, G., Archer, D.B., and Williamson, M.P. (1996). Solution structure of the granular starch binding domain of glucoamylase from Aspergillus niger by nuclear magnetic resonance spectroscopy. J Mol Biol *259*, 970-987.
147. Janecek, S., and Sevcik, J. (1999). The evolution of starch-binding domain. FEBS Lett *456*, 119-125.
148. Sorimachi, K., Le Gal-Coeffet, M.F., Williamson, G., Archer, D.B., and Williamson, M.P. (1997). Solution structure of the granular starch binding domain of Aspergillus niger glucoamylase bound to beta-cyclodextrin. Structure *5*, 647-661.
149. Knegtel, R.M., Strokopytov, B., Penninga, D., Faber, O.G., Rozeboom, H.J., Kalk, K.H., Dijkhuizen, L., and Dijkstra, B.W. (1995). Crystallographic studies of the interaction of cyclodextrin glycosyltransferase from Bacillus circulans strain 251 with natural substrates and products. J Biol Chem *270*, 29256-29264.

150. Larson, S.B., Greenwood, A., Cascio, D., Day, J., and McPherson, A. (1994). Refined molecular structure of pig pancreatic alpha-amylase at 2.1 A resolution. J Mol Biol *235*, 1560-1584.
151. Mikami, B., Hehre, E.J., Sato, M., Katsube, Y., Hirose, M., Morita, Y., and Sacchettini, J.C. (1993). The 2.0-A resolution structure of soybean beta-amylase complexed with alpha-cyclodextrin. Biochemistry *32*, 6836-6845.
152. Svensson, B., Jespersen, H., Sierks, M.R., and MacGregor, E.A. (1989). Sequence homology between putative raw-starch binding domains from different starch-degrading enzymes. Biochem J *264*, 309-311.
153. Williamson, M.P., Le Gal-Coeffet, M.F., Sorimachi, K., Furniss, C.S., Archer, D.B., and Williamson, G. (1997). Function of conserved tryptophans in the Aspergillus niger glucoamylase 1 starch binding domain. Biochemistry *36*, 7535-7539.
154. Penninga, D., van der Veen, B.A., Knegtel, R.M., van Hijum, S.A., Rozeboom, H.J., Kalk, K.H., Dijkstra, B.W., and Dijkhuizen, L. (1996). The raw starch binding domain of cyclodextrin glycosyltransferase from Bacillus circulans strain 251. J Biol Chem *271*, 32777-32784.
155. Liu, Y.N., Lai, Y.T., Chou, W.I., Chang, M.D., and Lyu, P.C. (2007). Solution structure of family 21 carbohydrate-binding module from Rhizopus oryzae glucoamylase. Biochem J *403*, 21-30.
156. Polekhina, G., Gupta, A., van Denderen, B.J., Feil, S.C., Kemp, B.E., Stapleton, D., and Parker, M.W. (2005). Structural basis for glycogen recognition by AMP-activated protein kinase. Structure *13*, 1453-1462.
157. Kotting, O., Santelia, D., Edner, C., Eicke, S., Marthaler, T., Gentry, M.S., Comparot-Moss, S., Chen, J., Smith, A.M., Steup, M., et al. (2009). STARCH-EXCESS4 is a laforin-like Phosphoglucan phosphatase required for starch degradation in Arabidopsis thaliana. Plant Cell *21*, 334-346.
158. Hejazi, M., Fettke, J., Kotting, O., Zeeman, S.C., and Steup, M. (2010). The Laforin-like dual-specificity phosphatase SEX4 from Arabidopsis hydrolyzes both C6- and C3-phosphate esters introduced by starch-related dikinases and thereby affects phase transition of alpha-glucans. Plant Physiol *152*, 711-722.
159. Lammerts van Bueren, A., Finn, R., Ausio, J., and Boraston, A.B. (2004). Alpha-glucan recognition by a new family of carbohydrate-binding modules found primarily in bacterial pathogens. Biochemistry *43*, 15633-15642.
160. Mikami, B., Iwamoto, H., Malle, D., Yoon, H.J., Demirkan-Sarikaya, E., Mezaki, Y., and Katsuya, Y. (2006). Crystal structure of pullulanase: evidence for parallel binding of oligosaccharides in the active site. J Mol Biol *359*, 690-707.
161. Bouju, S., Lignon, M.F., Pietu, G., Le Cunff, M., Leger, J.J., Auffray, C., and Dechesne, C.A. (1998). Molecular cloning and functional expression of a novel human gene encoding two 41-43 kDa skeletal muscle internal membrane proteins. Biochem J *335 (Pt 3)*, 549-556.

162. Pietu, G., Alibert, O., Guichard, V., Lamy, B., Bois, F., Leroy, E., Mariage-Sampson, R., Houlgatte, R., Soularue, P., and Auffray, C. (1996). Novel gene transcripts preferentially expressed in human muscles revealed by quantitative hybridization of a high density cDNA array. Genome Res *6*, 492-503.
163. Janecek, S. (2002). A motif of a microbial starch-binding domain found in human genethonin. Bioinformatics *18*, 1534-1537.
164. Minassian, B.A., Ianzano, L., Meloche, M., Andermann, E., Rouleau, G.A., Delgado-Escueta, A.V., and Scherer, S.W. (2000). Mutation spectrum and predicted function of laforin in Lafora's progressive myoclonus epilepsy. Neurology *55*, 341-346.
165. Giardina, T., Gunning, A.P., Juge, N., Faulds, C.B., Furniss, C.S., Svensson, B., Morris, V.J., and Williamson, G. (2001). Both binding sites of the starch-binding domain of Aspergillus niger glucoamylase are essential for inducing a conformational change in amylose. J Mol Biol *313*, 1149-1159.
166. Stapleton, D., Nelson, C., Parsawar, K., McClain, D., Gilbert-Wilson, R., Barker, E., Rudd, B., Brown, K., Hendrix, W., O'Donnell, P., et al. (2010). Analysis of hepatic glycogen-associated proteins. Proteomics *10*, 2320-2329.
167. Behrends, C., Sowa, M.E., Gygi, S.P., and Harper, J.W. (2010). Network organization of the human autophagy system. Nature *466*, 68-76.
168. Jiang, S., Heller, B., Tagliabracci, V.S., Zhai, L., Irimia, J.M., DePaoli-Roach, A.A., Wells, C.D., Skurat, A.V., and Roach, P.J. (2010). Starch binding domain-containing protein 1/genethonin 1 is a novel participant in glycogen metabolism. J Biol Chem *285*, 34960-34971.
169. Clark, S.J. (1957). Cellular differentiation in the kidneys of newborn mice studied with the electron microscope. J Biophys Biochem Cytol. *3*, 349-362.
170. De Duve, C., and Wattiaux, R. (1966). Functions of lysosomes. Annu Rev Physiol *28*, 435-492.
171. Yin, X.M., Ding, W.X., and Gao, W. (2008). Autophagy in the liver. Hepatology *47*, 1773-1785.
172. Tanida, I. (2011). Autophagy basics. Microbiol Immunol *55*, 1-11.
173. Klionsky, D.J. (2007). Autophagy: from phenomenology to molecular understanding in less than a decade. Nat Rev Mol Cell Biol *8*, 931-937.
174. Cuervo, A.M. (2004). Autophagy: in sickness and in health. Trends Cell Biol *14*, 70-77.
175. Klionsky, D.J., and Emr, S.D. (2000). Autophagy as a regulated pathway of cellular degradation. Science *290*, 1717-1721.
176. Stromhaug, P.E., Berg, T.O., Fengsrud, M., and Seglen, P.O. (1998). Purification and characterization of autophagosomes from rat hepatocytes. Biochem J *335 (Pt 2)*, 217-224.
177. Ohsumi, Y. (2001). Molecular dissection of autophagy: two ubiquitin-like systems. Nat Rev Mol Cell Biol *2*, 211-216.

178. Mizushima, N., Sugita, H., Yoshimori, T., and Ohsumi, Y. (1998). A new protein conjugation system in human. The counterpart of the yeast Apg12p conjugation system essential for autophagy. The Journal of biological chemistry *273*, 33889-33892.
179. Geng, J., and Klionsky, D.J. (2008). The Atg8 and Atg12 ubiquitin-like conjugation systems in macroautophagy. 'Protein modifications: beyond the usual suspects' review series. EMBO Rep *9*, 859-864.
180. Mizushima, N., Noda, T., Yoshimori, T., Tanaka, Y., Ishii, T., George, M.D., Klionsky, D.J., Ohsumi, M., and Ohsumi, Y. (1998). A protein conjugation system essential for autophagy. Nature *395*, 395-398.
181. Shintani, T., Mizushima, N., Ogawa, Y., Matsuura, A., Noda, T., and Ohsumi, Y. (1999). Apg10p, a novel protein-conjugating enzyme essential for autophagy in yeast. EMBO J *18*, 5234-5241.
182. Mizushima, N., Noda, T., and Ohsumi, Y. (1999). Apg16p is required for the function of the Apg12p-Apg5p conjugate in the yeast autophagy pathway. EMBO J *18*, 3888-3896.
183. Ichimura, Y., Kirisako, T., Takao, T., Satomi, Y., Shimonishi, Y., Ishihara, N., Mizushima, N., Tanida, I., Kominami, E., Ohsumi, M., et al. (2000). A ubiquitin-like system mediates protein lipidation. Nature *408*, 488-492.
184. Suzuki, K., Kubota, Y., Sekito, T., and Ohsumi, Y. (2007). Hierarchy of Atg proteins in pre-autophagosomal structure organization. Genes Cells *12*, 209-218.
185. Hanada, T., Noda, N.N., Satomi, Y., Ichimura, Y., Fujioka, Y., Takao, T., Inagaki, F., and Ohsumi, Y. (2007). The Atg12-Atg5 conjugate has a novel E3-like activity for protein lipidation in autophagy. The Journal of biological chemistry *282*, 37298-37302.
186. Kirisako, T., Ichimura, Y., Okada, H., Kabeya, Y., Mizushima, N., Yoshimori, T., Ohsumi, M., Takao, T., Noda, T., and Ohsumi, Y. (2000). The reversible modification regulates the membrane-binding state of Apg8/Aut7 essential for autophagy and the cytoplasm to vacuole targeting pathway. J Cell Biol *151*, 263-276.
187. Levine, B., and Klionsky, D.J. (2004). Development by self-digestion: molecular mechanisms and biological functions of autophagy. Dev Cell *6*, 463-477.
188. Qu, X., Zou, Z., Sun, Q., Luby-Phelps, K., Cheng, P., Hogan, R.N., Gilpin, C., and Levine, B. (2007). Autophagy gene-dependent clearance of apoptotic cells during embryonic development. Cell *128*, 931-946.
189. Tanida, I., Wakabayashi, M., Kanematsu, T., Minematsu-Ikeguchi, N., Sou, Y.S., Hirata, M., Ueno, T., and Kominami, E. (2006). Lysosomal turnover of GABARAP-phospholipid conjugate is activated during differentiation of C2C12 cells to myotubes without inactivation of the mTor kinase-signaling pathway. Autophagy *2*, 264-271.
190. Nishida, Y., Arakawa, S., Fujitani, K., Yamaguchi, H., Mizuta, T., Kanaseki, T., Komatsu, M., Otsu, K., Tsujimoto, Y., and Shimizu, S. (2009). Discovery of Atg5/Atg7-independent alternative macroautophagy. Nature *461*, 654-658.

191. Gutierrez, M.G., Master, S.S., Singh, S.B., Taylor, G.A., Colombo, M.I., and Deretic, V. (2004). Autophagy is a defense mechanism inhibiting BCG and Mycobacterium tuberculosis survival in infected macrophages. Cell *119*, 753-766.
192. Kon, M., and Cuervo, A.M. (2010). Chaperone-mediated autophagy in health and disease. FEBS Lett *584*, 1399-1404.
193. Li, W., Yang, Q., and Mao, Z. (2011). Chaperone-mediated autophagy: machinery, regulation and biological consequences. Cell Mol Life Sci *68*, 749-763.
194. Johansen, T., and Lamark, T. (2011). Selective autophagy mediated by autophagic adapter proteins. Autophagy *7*, 1-18.
195. Lynch-Day, M.A., and Klionsky, D.J. (2010). The Cvt pathway as a model for selective autophagy. FEBS Lett *584*, 1359-1366.
196. Komatsu, M., and Ichimura, Y. (2010). Selective autophagy regulates various cellular functions. Genes Cells *15*, 923-933.
197. Rabinowitz, J.D., and White, E. (2010). Autophagy and metabolism. Science *330*, 1344-1348.
198. Kirkin, V., Lamark, T., Sou, Y.S., Bjorkoy, G., Nunn, J.L., Bruun, J.A., Shvets, E., McEwan, D.G., Clausen, T.H., Wild, P., et al. (2009). A role for NBR1 in autophagosomal degradation of ubiquitinated substrates. Mol Cell *33*, 505-516.
199. Novak, I., Kirkin, V., McEwan, D.G., Zhang, J., Wild, P., Rozenknop, A., Rogov, V., Lohr, F., Popovic, D., Occhipinti, A., et al. (2010). Nix is a selective autophagy receptor for mitochondrial clearance. EMBO Rep *11*, 45-51.
200. Weidberg, H., Shvets, E., Shpilka, T., Shimron, F., Shinder, V., and Elazar, Z. (2010). LC3 and GATE-16/GABARAP subfamilies are both essential yet act differently in autophagosome biogenesis. EMBO J *29*, 1792-1802.
201. Hemelaar, J., Lelyveld, V.S., Kessler, B.M., and Ploegh, H.L. (2003). A single protease, Apg4B, is specific for the autophagy-related ubiquitin-like proteins GATE-16, MAP1-LC3, GABARAP, and Apg8L. The Journal of biological chemistry *278*, 51841-51850.
202. Tanida, I., Ueno, T., and Kominami, E. (2008). LC3 and Autophagy. Methods Mol Biol *445*, 77-88.
203. Leidenheimer, N.J., Browning, M.D., and Harris, R.A. (1991). GABAA receptor phosphorylation: multiple sites, actions and artifacts. Trends Pharmacol Sci *12*, 84-87.
204. Moss, S.J., Smart, T.G., Blackstone, C.D., and Huganir, R.L. (1992). Functional modulation of GABAA receptors by cAMP-dependent protein phosphorylation. Science *257*, 661-665.
205. Nusser, Z., Roberts, J.D., Baude, A., Richards, J.G., and Somogyi, P. (1995). Relative densities of synaptic and extrasynaptic GABAA receptors on cerebellar granule cells as determined by a quantitative immunogold method. J Neurosci *15*, 2948-2960.

206. Wang, H., Bedford, F.K., Brandon, N.J., Moss, S.J., and Olsen, R.W. (1999). GABA(A)-receptor-associated protein links GABA(A) receptors and the cytoskeleton. Nature *397*, 69-72.
207. Kabeya, Y., Mizushima, N., Yamamoto, A., Oshitani-Okamoto, S., Ohsumi, Y., and Yoshimori, T. (2004). LC3, GABARAP and GATE16 localize to autophagosomal membrane depending on form-II formation. J Cell Sci *117*, 2805-2812.
208. Pellerin, I., Vuillermoz, C., Jouvenot, M., Ordener, C., Royez, M., and Adessi, G.L. (1993). Identification and characterization of an early estrogen-regulated RNA in cultured guinea-pig endometrial cells. Mol Cell Endocrinol *90*, R17-21.
209. Vernier-Magnin, S., Muller, S., Sallot, M., Radom, J., Musard, J.F., Adami, P., Dulieu, P., Remy-Martin, J.P., Jouvenot, M., and Fraichard, A. (2001). A novel early estrogen-regulated gene gec1 encodes a protein related to GABARAP. Biochem Biophys Res Commun *284*, 118-125.
210. Xin, Y., Yu, L., Chen, Z., Zheng, L., Fu, Q., Jiang, J., Zhang, P., Gong, R., and Zhao, S. (2001). Cloning, expression patterns, and chromosome localization of three human and two mouse homologues of GABA(A) receptor-associated protein. Genomics *74*, 408-413.
211. Chakrama, F.Z., Seguin-Py, S., Le Grand, J.N., Fraichard, A., Delage-Mourroux, R., Despouy, G., Perez, V., Jouvenot, M., and Boyer-Guittaut, M. (2010). GABARAPL1 (GEC1) associates with autophagic vesicles. Autophagy *6*.
212. Sagiv, Y., Legesse-Miller, A., Porat, A., and Elazar, Z. (2000). GATE-16, a membrane transport modulator, interacts with NSF and the Golgi v-SNARE GOS-28. EMBO J *19*, 1494-1504.
213. Noda, N.N., Ohsumi, Y., and Inagaki, F. (2010). Atg8-family interacting motif crucial for selective autophagy. FEBS Lett *584*, 1379-1385.
214. Noda, N.N., Kumeta, H., Nakatogawa, H., Satoo, K., Adachi, W., Ishii, J., Fujioka, Y., Ohsumi, Y., and Inagaki, F. (2008). Structural basis of target recognition by Atg8/LC3 during selective autophagy. Genes Cells *13*, 1211-1218.
215. Platt, T., Muller-Hill, B and Miller, J. H. (1972). Assay of Beta-galactosidase. In Experiments in molecular biology., J.H. Miller, ed. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 352-355.
216. Tagliabracci, V.S., Turnbull, J., Wang, W., Girard, J.M., Zhao, X., Skurat, A.V., Delgado-Escueta, A.V., Minassian, B.A., Depaoli-Roach, A.A., and Roach, P.J. (2007). Laforin is a glycogen phosphatase, deficiency of which leads to elevated phosphorylation of glycogen in vivo. Proc Natl Acad Sci U S A *104*, 19262-19266.
217. Skurat, A.V., Dietrich, A.D., and Roach, P.J. (2000). Glycogen synthase sensitivity to insulin and glucose-6-phosphate is mediated by both NH2- and COOH-terminal phosphorylation sites. Diabetes *49*, 1096-1100.

218. Pederson, B.A., Cope, C.R., Schroeder, J.M., Smith, M.W., Irimia, J.M., Thurberg, B.L., DePaoli-Roach, A.A., and Roach, P.J. (2005). Exercise capacity of mice genetically lacking muscle glycogen synthase: in mice, muscle glycogen is not essential for exercise. J Biol Chem *280*, 17260-17265.
219. Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem *72*, 248-254.
220. Schaart, G., Hesselink, R.P., Keizer, H.A., van Kranenburg, G., Drost, M.R., and Hesselink, M.K. (2004). A modified PAS stain combined with immunofluorescence for quantitative analyses of glycogen in muscle sections. Histochem Cell Biol *122*, 161-169.
221. Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res *29*, e45.
222. Blennow, A., Nielsen, T.H., Baunsgaard, L., Mikkelsen, R., and Engelsen, S.B. (2002). Starch phosphorylation: a new front line in starch research. Trends Plant Sci *7*, 445-450.
223. Sakai, M., Austin, J., Witmer, F., and Trueb, L. (1970). Studies in myoclonus epilepsy (Lafora body form). II. Polyglucosans in the systemic deposits of myoclonus epilepsy and in corpora amylacea. Neurology *20*, 160-176.
224. Breslow, J.L., Sloan, H.R., Ferrans, V.J., Anderson, J.L., and Levy, R.I. (1973). Characterization of the mouse liver cell line FL83B. Exp Cell Res *78*, 441-453.
225. Wang, J., Stuckey, J.A., Wishart, M.J., and Dixon, J.E. (2002). A unique carbohydrate binding domain targets the lafora disease phosphatase to glycogen. J Biol Chem *277*, 2377-2380.
226. Wang, W., and Roach, P.J. (2004). Glycogen and related polysaccharides inhibit the laforin dual-specificity protein phosphatase. Biochem Biophys Res Commun *325*, 726-730.
227. Romero, P., Obradovic, Z., and Dunker, A.K. (2004). Natively disordered proteins: functions and predictions. Appl Bioinformatics *3*, 105-113.
228. Wang, W., Parker, G.E., Skurat, A.V., Raben, N., DePaoli-Roach, A.A., and Roach, P.J. (2006). Relationship between glycogen accumulation and the laforin dual specificity phosphatase. Biochem Biophys Res Commun *350*, 588-592.
229. Fukuda, T., Roberts, A., Ahearn, M., Zaal, K., Ralston, E., Plotz, P.H., and Raben, N. (2006). Autophagy and lysosomes in Pompe disease. Autophagy *2*, 318-320.
230. Schiaffino, S., and Hanzlikova, V. (1972). Autophagic degradation of glycogen in skeletal muscles of the newborn rat. J Cell Biol *52*, 41-51.
231. Kondomerkos, D.J., Kalamidas, S.A., Kotoulas, O.B., and Hann, A.C. (2005). Glycogen autophagy in the liver and heart of newborn rats. The effects of glucagon, adrenalin or rapamycin. Histol Histopathol *20*, 689-696.

232. DiMauro, S., and Lamperti, C. (2001). Muscle glycogenoses. Muscle Nerve *24*, 984-999.
233. Hu, Z.Z., Valencia, J.C., Huang, H., Chi, A., Shabanowitz, J., Hearing, V.J., Appella, E., and Wu, C. (2007). Comparative Bioinformatics Analyses and Profiling of Lysosome-Related Organelle Proteomes. Int J Mass Spectrom *259*, 147-160.

CURRICULUM VITAE

Sixin Jiang

EDUCATION

2011 Ph.D., Biochemistry and Molecular Biology, Indiana University
2004 M.S., Microbial genetics, Institute of Microbiology, Chinese Academy of Sciences
2000 B.S., Plant protection, Nanjing Agricultural University

RESEARCH EXPERIENCE

2004 – 2011 Department of Biochemistry and Molecular Biology, Indiana University School of Medicine
Glycogen metabolism.
Dissertation: Starch binding domain containing protein 1: a novel participant in glycogen metabolism.

2001 – 2004 Institute of Microbiology, Chinese Academy of Sciences
Molecular Genetics and Breeding of Yeast.
MS thesis: Cloning and co-expression of malolactic fermentation (MLF) related genes from *Oenococcus oeni* in *Saccharomyces cerevisiae* and the effect on recombinant *S. cerevisiae* metabolism.

1999 Institute of Plant Protection, Chinese Academy of Agricultural Sciences
Maize Dwarf Mosaic Virus (MDMV).
Project: Correlation between corn species and resistance to MDMV and comparison of MDMV content in different parts of the corn plant.

PUBLICATIONS

Jiang S, Heller B, Tagliabracci VS, Zhai L, Irimia JM, Depaoli-Roach AA, Wells CD, Skurat AV, Roach PJ. Starch binding domain containing protein 1/genethonin 1 is a novel participant in glycogen metabolism. (2010) J Biol Chem. 285: 34960-34971.

Jiang S, Liu Y, He, X, Guo X, Zhang B. Cloning of *mleP* gene from *Oenococcus oeni* and Expression in *Saccharomyces cerevisiae*. (2004) Wei Sheng Wu Xue Bao. 44: 465-468. (Chinese)

Liu, Y, Jiang S, Li H, Zhang B. Cloning and Expression of MLF-related Genes in Yeast. (2003) Sheng Wu Gong Cheng Xue Bao. 23: 27-30. (Chinese)

Liu Y, Li H, Jiang S, Zhang B. Advances of Research on Enzyme and Genes of Malolactic Fermentation. (2003) Wei Sheng Wu Xue Tong Bao. 30: 103-107. (Chinese)

Liu C, He X, Jiang S**,** Qu N, Zhang B. Breeding of Excellent Baker's Yeast Strain with Good Flocculation. (2003) Wei Sheng Wu Xue Bao. 43: 659-665. (Chinese)

POSTERS AND PRESENTATIONS

Starch binding domain containing protein 1 (Stbd1): A Novel Participant in Glycogen Metabolism. (2010) Sigma Xi Graduate Research Competition. Indiana University Medical Center Chapter of Sigma Xi, Indianapolis, IN.

The role of the novel polysaccharide binding protein Genethonin 1 in glycogen metabolism. (2008) Biochemistry Research Day. Indiana University School of Medicine, Indianpolis, IN.

Interactions between glycogen and hexosamine biosynthetic pathways. (2007) Biochemistry Research Day. Indiana University School of Medicine, Indianapolis, IN

AWARDS

Second Prize of Institute Scholarship, 2002-2003
Honor Graduate in Department of Plant Protection, 2000
First Prize of National Excellent Students Scholarship, 1998
Red Sun Scholarship, 1998

PROFESSIONAL AFFILIATION

American Association for the Advancement of Science, 2009-present